Contents

Chapter 6. Data Processing

(Page numbers here omit the chapter prefix, 6-, which appears as part of the number of individual pages and appendixes. Each chapter is paginated separately for convenience of issue)

	Page
INTRODUCTION	5
ORGANIZATIONAL STRUCTURE	5
ESTABLISHING THE PROCESSING OFFICES	5
General Information	5
Planning for and Obtaining the Sites	6
Opening the Processing Offices	6
Logistics	6
Supplies	6
Staffing	7
Training	7
Payrolling	9
Security	9
Management Information System	10
RECEIPT AND CHECK-IN	10
General Procedures	10
Quality control	12
Problems encountered	12
Problem resolution	12
Geographic Problem Resolution	13
CONTROL OF MATERIALS	13
Introduction	13
Automated Inventory Control System (AICS)	13
Introduction	13
Equipment	14
General procedures	14
AICS bar-code station operations	14
Bar-code label printing	15
Backup Manual-Control Operations	15
Libraries	15
Introduction	15
Miscellaneous materials library	15
ED questionnaire library	16
Address register library	16
Reference materials library	16

100-PERCENT PRETABULATION PROCESSING	17
Introduction	17
100-Percent Microfilming Operations	17
General information	17
Camera preparation	17
Camera operations	18
Camera maintenance	19
Film processing	19
100-Percent FOSDIC Processing	19
General information	19
Staffing and training	20
FOSDIC operations	20
FOSDIC maintenance	20
Film Library	20
Diary Review	21
General information	21
Procedures	21
Referrals	22
Quality review	22
Diary review closeout	23
SAMPLE PRETABULATION PROCESSING	23
General Information	23
Clerical Coding of Handwritten Responses	24
Constraints on the sample coding operations	24
Quality Control	24
Microfilming of Coded Questionnaires	24
Film Library	25
Sample FOSDIC Processing Operations	25
Sample Diary Review	26
OTHER PROCESSING CENTER OPERATIONS	26
Automated Inventory and Control System (AICS) Operations	26
Label printing	26
Keying Diary Review and Other Programs	26
Diary review	26
Post-enumeration survey	26
Other programs	27
Count Review	27
CLOSING THE PROCESSING OFFICES	27
HEADQUARTERS PROCESSING OPERATIONS	27
General Procedures	27
Detail-File Preparation	28
Computer Edits	28
Sample Weighting	28
Analyzers	29
Confidentiality and Suppression	29
Data Review	30
TARIII ATION	30



Chapter 6. Data Processing

INTRODUCTION

As each district office closed, it boxed and shipped its questionnaires, address registers, and miscellaneous materials by truck to one of three clerical processing offices—in Jeffersonville, IN, New Orleans, LA, or Laguna Niguel, CA—where the data from about 90 million report forms would be coded where necessary, microfilmed, scanned, and the data transmitted to the Bureau's headquarters in Suitland, MD, for final processing and tabulation. The preliminary processing task required space for storage and warehousing, large clerical staffs, and a mass of electronic and mechanical equipment that could not be accommodated at a single location.

ORGANIZATIONAL STRUCTURE

The Bureau created a Decennial Processing Staff (DPS), under the Associate Director for Field Operations and his Assistant Director for Processing, to organize and control the clerical and precomputer processing of census questionnaires. The DPS was abolished in March 1982, after its task was completed. The Decennial Census Division (DCD), under the Associate Director for Demographic Fields and his Assistant Director for Demographic Censuses, helped plan and conduct the census effort. The Field Division (FLD) worked closely with DPS and DCD on the processing center operations. During diary review and computer processing, the two subject matter divisions, Population (POP) and Housing (HOUS), and several computer-oriented and administrative divisions were heavily involved in processing activities. The Organization and Management Systems Division set production standards, while the Statistical Methods Division established the standards for quality control.

Each processing office had a manager and two assistant managers—one for administration, the other for operations. The assistant manager for administration was responsible for providing general support and "housekeeping" services, and supervised administrative support (i.e. facilities and space management, local payroll operations, hiring, security, etc.), training, automated-data processing support, production standards and incentive awards (Laguna Niguel and New Orleans offices only), and information control units. The assistant manager for operations directed the processing operations themselves, supervising the coding and general processing operations units. These positions, and other top supervisory slots, were filled by career Census Bureau personnel.

In a shift from the 1970 census organizational pattern, when basically the same processing organization controlled both production and quality control (QC), each processing office in the 1980 census had a separate QC and evaluation staff reporting to the office manager. This staff implemented and managed QC and evaluation programs for all processing activities within the office; and hired, scheduled or trained, and managed all personnel engaged in carrying out the QC and evaluation (as distinguished from production) requirements. The staff included a unit that handled all documents affecting its operations.

ESTABLISHING THE PROCESSING OFFICES General Information

While the three processing offices opened well before the census and took part in a variety of precensus activities (see chs. 3 and 5 for details), their principal job was to process the census questionnaires and other materials and transfer the data contained therein to computer tape for processing and tabulation. They performed the following major operations:

- Receiving and unloading trucks arriving from district offices, and sorting, palletizing, and holding the materials received until they could be routed to check-in.
- Checking in questionnaires, address registers, and miscellanenous materials to establish an initial inventory of district office receipts. (For boxes of questionnaires and address registers, the check-in included affixing a bar-code label to assist in tracking these materials through the processing system.)
- 3. Shelving materials received in the appropriate libraries to provide a permanent location for all materials to be kept while they were not in the processing flow.
- 4. Controlling the questionnaires through the basic processing operations required to convert the data contained to a computer-usable form. These operations were: 100-percent microfilming, 100-percent diary review, 100-percent remicrofilming, sample coding, sample microfilming, sample diary review, and sample remicrofilming.
- 5. Resolving geographic problems identified in check-in and diary review.
- 6. Conducting evaluations and special studies of census content and methodology (see ch. 9).
- Implementing a quality-control (QC) program for all activities performed in the processing offices.

Planning for and Obtaining the Sites

Early in the 1980 census planning period, a committee was formed to develop clerical processing office procedures to use in transferring the data from the questionnaires onto computer tapes for tabulation and cross classification. This expensive, lengthy, and labor-intensive task had taxed the one site used for the 1970 census, so the Bureau decided to distribute the 1980 clerical processing among four strategic locations-later changed to three—Jeffersonville, IN, New Orleans, LA, and Laguna Niguel, CA.

At Jeffersonville, the Census Bureau established the processing office within the Data Preparation Division (DPD) office already in place, as had been done for 1970. The Laguna Niguel and New Orleans sites were selected in part because (1) there were already large Federal Government office facilities in operation; the General Services Administration (GSA) managed such a complex in Laguna Niguel and the National Aeronautics and Space Administration (NASA) had offices in New Orleans, from which floor space (approximately 320,000 square feet was needed for each processing office) and equipment could be leased; (2) the respective areas had transportation and communications services to adequately service each office; and (3) the labor pool available could provide the necessary work force.

Opening the Processing Offices

The New Orleans and Laguna Niguel offices opened in February, and the Jeffersonville processing operation in August, 1978, to accommodate certain precensus operations, such as kit assembly and tape-address register/advance post-office check (TAR/APOC) and prelist address register keying (see chs. 5 and 3 respectively for details). The early opening also made it possible to prepare for a timely handling of the district office output of questionnaires and address registers.

Logistics

Supplies—The Decennial Processing Staff (DPS) at Bureau headquarters procured such key items as the following:

Item	Number
Cardboard desks	6,670
Rolling bins	6,216
Plastic inserts	800,000
Steel shelving	182,2751

¹These were steel shelves for boxes of questionnaires and address registers (AR's), and maps. The shelves for questionnaire boxes were 30" deep by 42" wide and were assembled into units with 10" between shelves, allowing for three boxes to be stored side by side, and stacked four deep. The AR boxes were 16" deep by 11 1/2" wide, and 1/2" high. The AR shelving was assembled with 13" between shelves, and the AR's were stacked on the shelves on their sides. The shelves could be assembled into units of varying overall height to make maximum use of available space.

At the same time, DPS and Field Division arranged, through the 12 regional census centers (RCC's), for shipment of unneeded furniture and other supplies from some of the 409 district offices to the processing offices. The desks were for employees' work stations, while the rolling bins were used to move materials (boxes of questionnaires, address registers, etc.). The plastic inserts were for all enumeration district (ED) questionnaire boxes and the steel shelving was used for storage of processing materials, including questionnaires and other documents.

The procurement, rental, and distribution procedures used for the 1980 census equipment and supplies were similar to those employed for 1970. DPS procured and/or rented most equipment and supplies for the processing offices through the Administrative Services Division (ASD), which forwarded requests for equipment and supplies to GSA or private vendors as appropriate. Some specialized equipment for processing offices required early procurement, with delivery dates scheduled at the completion of site preparation. The Department of Commerce (DOC) purchasing staff had to approve procurement of specialized items; i.e., automated data-processing equipment, film processing equipment, steel shelving, and so on.

For 1980, DPS identified the equipment and supplies needed for the processing offices—including all printed materials used exclusively in the processing sites, such as evaluation forms and procedures, and Population Division (POP) provided all coding manuals and training materials. With five different units ordering and reordering supplies, there was considerable difficulty in controlling delivery and receipt dates for certain supplies, and the Bureau experienced some processing delays as a result.

Supplies, printed matter, equipment, and furniture requirements were based on estimates of operational needs, and until operational procedures had been established, these had to be calculated from previous censuses as updated by test censuses. A great deal of planning went into the development of comprehensive lists of equipment, furniture, and supplies required in taking and processing the census. Lists of materials were circulated to various segments of the Bureau for comments, and conferences were held to decide on final requirements. Comprehensive planning eliminated most requirements for emergency purchases, but some problems did arise and had to be dealt with immediately.

For example:

- Steel shelving units had been received in the New Orleans processing office without sufficient quantities of cross braces, nuts, and bolts.
- The roof collapsed on building 220 of the New Orleans processing office, damaging or destroying many expendable supplies and steel shelving.
- Occasionally, items that had been ordered did not arrive in time for an operation.

When emergency purchasing requirements arose in the processing offices, they were handled in one of the following ways:

- Purchase order. When time allowed, a form CD-45 (Supply, Equipment, or Service Order) was prepared and forwarded to the headquarters procurement office, ASD. To expedite delivery in emergency situations, ASD telephoned purchase order numbers to a vendor.
- Emergency procurement authority. Authority was delegated to the chiefs of the New Orleans and Laguna Niguel processing offices for the emergency purchase of noncapital items and services, not to exceed \$500 per order. Invoices for payment were approved by the processing office chiefs and forwarded to Finance Division (FIN).
- 3. Shipment from other offices. When emergency requirements arose (particularly in the case of specialized items such as white correction dots and colored tape), supplies were sent from headquarters, DO's, or processing offices to cover the immediate need until formal orders could be processed.
- 4. Imprest fund. The New Orleans and Laguna Niguel processing offices were allowed to establish \$1,000 imprest funds.
- 5. Blanket purchase orders. The processing offices issued these for supplies such as lumber, plywood, cardboard cartons, envelopes, hardware, and office supplies, and drew on them when emergency requirements arose.
- 6. Printing. The processing offices were authorized to deal directly with their respective (GPO) Government Printing Office facilities for emergency printing requirements.

Staffing—The Census Bureau took 18 months and employed approximately 6,300 temporary staff members at peak levels to complete the postenumeration processing at the three offices. The Decennial Processing Staff (DPS) at Bureau headquarters was created to organize and control the clerical and precomputer work at the processing sites, while the computer processing was done at the Suitland, MD headquarters. Top supervisors were selected primarily from among the Bureau's career staff; clerical employees were chosen through the civil service system or through special waivers when necessary. Professional and technical subject-matter review of the FOSDIC generated computer totals and the computer processing activity itself was done by headquarters staff.

The processing office staffs were recruited from the local labor force, and were to have a race-ethnic mix comparable to that of the surrounding population; hence affirmative action hiring plans were developed for each site. Senior supervisory and technical positions—i.e., office managers and assistant managers, branch chiefs, certain staff administrative positions, and selected technical positions—were filled by permanent

Census Bureau employees assigned from headquarters or the regional offices. (For the Jeffersonville office, DPD staff filled most of the senior slots.) The remaining supervisory positions, as well as the mass of the clerical, service, and other jobs, were filled from the local workforce.

The Census Bureau recruited personnel directly, as well as accepting referrals from various local civic, civil rights, and other appropriate sources. Each applicant had to meet the following requirements for possible hiring:

- Be able to do most office tasks and, for some positions, be capable of specified physical labor, such as loading or unloading pallets, lifting specified weights, moving rolling bins, and so on.
- Be at least 18 years of age, although 16- or 17-year olds could be hired if they met employment conditions set by local and State governments and were high school graduates or had equivalent education or work experience.
- 3. Be available to work a 40-hour work week and, for some jobs, to work specified shift schedules.
- 4. Anyone barred from a civil service examination could not be considered for employment.
- 5. Federal civil service annuitants were discouraged from applying.
- 6. A satisfactory work record prior to hiring. Poor job performance, immoral conduct, or conviction of a law violation since age 18 for something other than a minor traffic violation could be the basis for disqualification.

The principal means of selection for hiring was a written test administered to applicants, although this requirement was waived in some instances. Final selections for hiring were made by the personnel office at each processing site.

Recruiting, testing, and hiring began before the offices opened in 1978, and continued through much of their 4-year existence. With the general turnover of the workforce, over 11,000 people were hired.

Training—The training for over 200 different processing operations had to be conducted on parallel tracks at each of the three sites. Census requirements demanded that each employee be trained using a program that taught the same procedures in the same way in every session at each location in an extremely short time span. Since highly complex and sophisticated procedures had to be used to track literally millions of items, training techniques that simplified the instructions were essential. Because of the personnel attrition inherent in a large temporary operation, training had to be repeated over the life of the project. The training plan had to (1) minimize hiring a

permanent staff of "subject experts" to present each session, and (2) provide the right combination of instruction and practice, with the use of visual displays to shorten the initial instructional phase.

While planning the processing operation, the Census Bureau staff identified those processing operations (including administrative and support functions) requiring some form of class-room or on-the-job (OJT) training. Training for about 20 of these—the more technical- and/or action-intensive—required using video aids. For the remainder, the Bureau used other training media, such as OJT, special outside "vendor sponsored" programs, (as in film processing), and occasional class-room lecture techniques, using job aids—either procedures manuals or other reference material. The agency prepared training packages for each program that included a training guide, media presentation (if a video program), practice exercises or workbooks, and a job aid.

After analyzing the available training methodologies, the cost and time constraints, and the processing requirements, the staff selected videotaped modules prepared by headquarters staff as the primary training tools, rather than the verbatim guides followed in the field district offices. (Guides were used, however, in conjunction with audiovisual presentations.) An Office of Management and Budget (OMB) directive required that audiovisual materials be controlled, reviewed, and approved by the Commerce Audiovisual Review Committee before they were used officially. The Committee carried out its review in the first half of 1979, releasing the materials for use in June. An outside consulting firm assisted in planning the selection and subsequent production of the programs, and a joint agreement between the Bureau and the Office of Personnel Management (OPM) enabled the Bureau to use OPM's studios to produce these video programs at significant savings in time and money. The Bureau also developed a "Training Package Development Operating Plan" and a basic training guide as a precaution against late completion of the video materials.

The production schedules were tight for the video modules because of the limited availability of Government production studios. OPM required script writers, production managers, and "script doctors" to be used, but this proved time-consuming, causing significant scheduling problems. After experiencing considerable difficulty in preparing the video materials, the Census Bureau decided to use subject-matter and operational experts from the respective areas to design outlines for each training program. These outlines included visual aids (e.g., word cards, art work, and slides), and were used by the subject-matter experts to deliver the operational instructions directly onto videotape. DPS also contacted several other Government studios—at the Department of Health, Education, and Welfare (HEW), the National Institutes of Health (NIH), and the Department of Defense (DOD)—and obtained permission to use their facilities. DPS and OPM thus were producing video packages simultaneously.

The Census Bureau developed two "test" training programs for the dress rehearsal, using 21-inch color monitors and video recorders employing 3/4-inch tape. These proved successful, and for the census, the following equipment was procured and used:

- 1. 25-inch monitors—10 at each processing center.
- 2. 1/2-inch video players—six at each processing center.
- 3. 1/2-inch video recorders (to correct tapes and make additional tapes)—two at each processing center.
- 4. 1/2-inch portable audio tape recorders (to record the operations in session)—one at each processing center.
- Portable cameras (to make tapes in-house and for use with the portable recorders)—two each at Jeffersonville and New Orleans and one at Laguna Niguel.

Much of the training had to cover a variety of technical skills (for example, microfilming, library operations, and coding) for several different jobs, such as supervisor, control clerk, quality control clerk, and processing clerk. The instructions had certain elements in common, such as materials security, questionnaire condition, safety, and the overall processing system, and package designs employed a modular approach to minimize repetition in the training materials. Because each module was "customized" to address specific tasks within each of the operations, it was possible to construct specialized training sessions by assembling the right modules. This substantially reduced the time required for training by concentrating only on those areas the trainees were required to know about to do their jobs effectively. It also satisfied the consistency requirement, since each specialized session was universal throughout each office.

A single classification and control system was designed, with a series of module checklists for every operation that identified the modules trainees for a specific job would receive. (The instructional information common to all operations was combined within a single module and shown only once to each trainee.) Each module was given a classification number identifying it by processing operation (e.g., receipt), position (e.g., processing clerk), and type of material (e.g., audiovisual (A/V) tape).

The unique training requirements for the 1980 processing operation-training over 11,000 people during more than 500 separate sessions—called for close coordination of the training schedule. A training branch was created in each processing office, headed by a "master trainer" (education specialist), with a scheduling unit and a library unit, each headed by a training assistant. The scheduling unit received from the operational managers the training requests specifying the operational job type requiring training, verified available training space for the date requested, and determined what special arrangements, if any, had to be made. The library unit assembled the necessary modules and materials and notified the instructors to prepare

for the session. When each session was concluded, the library unit picked up any equipment and unused materials so that they would be available for other sessions. Training began in early June 1980. A total of 1,811 supervisors and over 9,000 clerks were selected and trained at the three processing office sites.

The training branches at the individual processing offices informed the operational staff of revisions to all materials maintained by the training library units. Minor problems encountered included discrepancies in the training system classification of materials (library number) and the volume-chapter-section identification system for the operational procedures manual. Most problems encountered were correctable and the existing instructions modifiable to alleviate special situations without serious consequences. Training for management and supervisors was formally scheduled, while the clerical training was scheduled on a "need" basis, depending on the starting dates of the operations and the requirements for replacement training and retraining.

All three processing offices had problems with regard to adequate space for training employees. At the Jeffersonville processing office, no permanent library space was available at first, so training materials were placed on rolling bins while the written plan for organizing the library was being followed. Once training began, some materials for several 100-percent operations arrived late, so the system had to be modified. After the initial management and supervisory training sessions were concluded and 100-percent operational training began, however, the system functioned as designed. At the New Orleans and Laguna Niguel offices, some of the assigned space was not suited for classroom training, reducing training effectiveness. The New Orleans office had further difficulties because shelving for the storage of training materials did not arrive on time and there were, at first, insufficient materials to support the number of trainees.

Overall, the biggest problems were lack of adequate space for both training and storage of training materials, lack of adequate training materials for some operations due to staff increases, late arrival of some materials for scheduled sessions, and the duplication and distribution of procedural revisions.

Payrolling—The processing office staffs were composed of "regular" Census Bureau permanent and temporary employees, and were paid on the standard civil service GS/GG scales for their respective grades. The respective processing offices' personnel units were responsible for timekeeping for all employees at their sites, but the actual payroll operation was handled through the payroll office at the Suitland headquarters. As the processing offices were activated and their staffs recruited, payroll clerks were added to the Suitland payroll office staff—one for each 500 employees at the processing offices.

When the processing office staffs reached their maximum strength, they employed approximately 6,300 people, and 13 extra payroll clerks on the headquarters staff handled pay records for that workforce.

The staffs were paid on the regular biweekly schedule, with checks issued at headquarters, and had the same options for automatic deposit or delivery of the checks as did other Census Bureau personnel. Checks to be issued directly were delivered to the processing sites by courier.

Security—The processing center operations required the temporary storage of massive amounts of confidential materials. Security was a major concern because of the risks of (1) destruction of address registers or of the original questionnaires before microfilming (the address registers were not microfilmed) and transmission of the data to headquarters, (2) violation of the confidentiality of the census records, and (3) damage to a variety of automated data processing, microfilming, and FOSDIC equipment, all vulnerable to fire or water damage. Each processing office designated a number of area security officers to assist in the overall security program.

Access to the processing offices was tightly controlled. There were uniformed guards at pedestrian entrances and exits, and standard Government fire control systems and damage control procedures were established. Each employee was issued a badge authorizing access to census offices, with additional badges worn by personnel requiring access to the questionnaire library and to the camera and automated data-processing areas. Visitors had to have special visitors' identification badges. All employees were reminded quarterly of their security and confidentiality responsibilities.

The security policy for the camera and automatic dataprocessing (ADP) areas required that such things as cameras, FOSDIC equipment, video training equipment, remote job entry (RJE) terminals, ENTREX data-entry equipment, and film processing equipment receive special protection from unauthorized use and malicious damage. Several steps were taken to implement this program:

- Each processing office appointed a permanent-staff member as ADP security coordinator.
- All entrances of the FOSDIC, RJE, ENTREX, and film processing areas were secured with cipher locks. Only persons whose jobs required access to this equipment had the cipher combination. The RJE rooms also were fitted with deadbolt locks, and rooms had to be locked when left unattended.
- An employee within the camera area was designated an area security officer to ensure that access to the camera room was limited to authorized personnel only.
- Visitors to the camera and ADP areas had to have the ADP security coordinator's permission and be escorted at all times.

- An ADP escort had to be present if these areas were to be cleaned after working hours.
- All camera and ADP areas had to be locked when unattended.

The processing offices used a number of terminals to enter the required data for the automated inventory and control system (AICS) and to guery these data. These terminals were located at key control points throughout the processing offices, but not always in areas that could be secured to meet minimum terminal and data security requirements. To prevent unauthorized use, each terminal had a lock interconnected with the unit's power supply switch. This required the use of a key before an operator could use an individual unit and permitted security control by unit regardless of location. The total communications network was disabled each evening by disconnecting the terminal communication lines at the network's communication concentrators. This automatically prevented access to the timesharing service. Finally, the acoustical couplers were stored in the locked terminal console cabinets each evening. The coupler provided an alterative means for terminal communications.

Management Information System

This automated system was designed for the 1980 census to provide pertinent information on the processing operations in a timely and efficient manner for use both in the offices and at headquarters. Every day, clerks keyed data received from the quality-assurance operations and grouped them into batches that were transmitted once a week to a main data base in Columbus, OH. From this base, each processing office printed out various types of output reports for its own use in management, and the Decennial Processing Staff's Inventory Control Branch distributed them weekly at Bureau headquarters.

The reports allowed management to make decisions about the coding and other processes and make comparisons among the three offices; for example, the summaries of coding production, error, number of enumeration districts processed, etc., showed both weekly and cumulative progress, while data on the distribution of production and error rates were useful in statistical testing.

The MIS included some programmed edits that checked for the validity of the data and for incompatible decision relationships (e.g., a decision to accept a unit of work but yet to retrain the coder). The MIS also did some simple linear regression analysis. In general, the system worked well after various problems were identified and corrected during testing. One major problem was solved by modifying the program to accept one set of data from one processing office at a time; prior to that, the computer in Columbus would only receive data if all three offices were transmitting the data at the same time. Other difficulties were detected later, such as delays in sending

records to the MIS keyers, errors in keying (such as identifying the wrong coder) despite 100-percent verification, and trouble weighting estimated error rates by computer when quality-control sampling rates were changed.

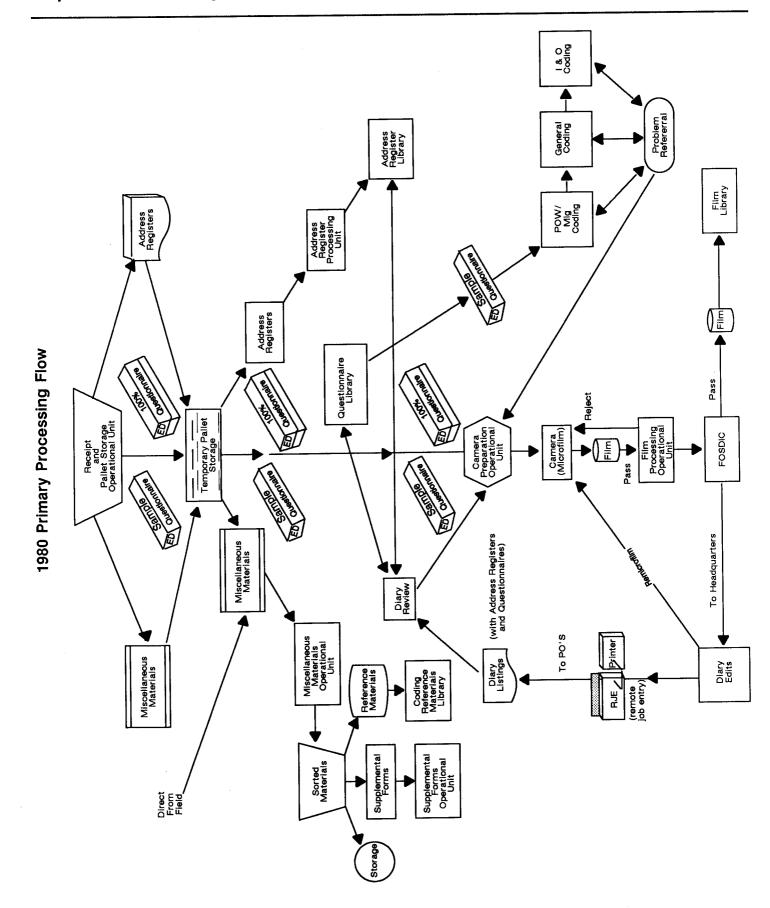
In an effort to eliminate indiscriminate changes, the Bureau instituted a document, form D-1440, Problem Identification and Resolution Record, which everyone perceiving a need for a change in procedures was required to use, even though solutions might be suggested and negotiated face-to-face or by telephone more or less simultaneously. The form D-1440, normally including a recommended solution, was sent to the Decennial Processing Staff (DPS), which circulated it to the appropriate divisions for review and return. DPS then disseminated the agreed-upon action. Several thousand changes were made in this way, normally with a 2-day turnaround.

RECEIPT AND CHECK-IN

General Procedures

The processing operation began with the 409 district offices shipping their ED boxes to one of the three processing offices. This was done on a flow basis from the first week of August 1980 until late December 1980, with a very large proportion of material arriving late in September (e.g., at the New Orleans office, almost 80 percent of receipts were received in the last 2 weeks of September). While some of the district offices (DO's) closed in August 1980, the majority of them remained open until mid September, and the last office closed in December 1980 (see ch. 5). The range of receipt dates for the 1980 decennial census was approximately 3 months later than the comparable dates for the 1970 decennial census, which delayed the start of the processing and forced the Census Bureau to compress the original 6-month processing schedule into only 3 1/2 months.

As materials arrived at the processing offices, they were sorted and loaded for bulk handling onto cargo pallets, which were placed in racks for temporary storage. These operations were subjected to formal quality assurance (QA) procedures. On request from the Information Control Branch (ICB), the Receipt and Pallet Storage Unit delivered pallets containing ED boxes, on a flow basis, by DO to the Check-In Unit. There the pallets were unloaded and the boxes of questionnaires stacked by ED on conveyor belts. (The conveyor belts were arranged like a horseshoe within the ED Check-In Unit.) The boxes were removed from the conveyor, by DO/ED, the bindings were cut, and the lids removed. The label of the top questionnaire, visible through a window in the packing material, was verified against the information listed on the external label for each ED box-i.e., the clerks determined that boxes marked as containing long forms did in fact contain long forms, and that the DO and ED numbers written on the box label were the same as the DO/ED on the questionnaire label. All damaged boxes, and



boxes in which DO/ED verification was not established, were sent to the Repair Unit for corrective action. The boxes of questionnaires for acceptable ED's were placed on a second conveyor for continued processing.

The appropriate bar code labels for each DO and ED were located and applied to the corresponding boxes. (The unique bar-code numbers were the primary identification device in the automated inventory and control system (AICS, see p. 13 below), and during processing were electronically scanned as the boxes were checked in and out of each work area. This made it possible, through the computer system, to tell where the materials for a given ED were at any time.) In 1970, one of the problems in the processing operation was the interchanging of box lids. The lids contained information on the processing status of the questionnaires in each box and identified the DO/ED involved. In large clerical operations, the lids were removed to process the questionnaires and often the wrong lid was placed on the box, resulting in questionnaires being lost or improperly processed. To prevent this from recurring, all identifying information for the 1980 questionnaires was placed on the end panel of the box bottom. The original box top was replaced with one that had a clear plastic insert in the end that covered the box-bottom panel. The bar-code label was placed so it could be read through a cutout or notched section of the plastic by a hand-held optical "wand" linked to the Bureau's AICS computer. If all the boxes were stacked in the same direction, verification of the DO/ED number could be made without having to handle each box manually.

Quality control—Quality-control checks were conducted throughout the entire operation to identify improperly labeled ED boxes, damaged plastic fronts, and similar problems. After the bar-code labels had been affixed to the ED boxes and the proper quality control review was completed, the ED boxes were placed on rolling bins, identified by ED range within DO, and bar-code read (by bin) into the AICS. ED's or ED boxes identified by the system as problem cases were removed from each bin and placed in a special bin for resolution. After the materials for problem ED's had been removed, the remaining boxes were sent to the ED Questionnaire Library.

The quality assurance (QA) clerks verified a sample of the ED's boxes. If problems were encountered, these materials were referred to the Problem Resolution Unit by way of the ED Check-In Unit control clerk. Acceptable ED boxes were passed to the end of the conveyor belt, loaded into designated range bins, and sent to the bar-code station for check-in to the ED library. The bins containing the original ED box lids were sent to the plastic-insert unit, where clerks used special box cutters that removed the end of the box lid and replaced it with a plastic insert. Filament tape was used to secure the plastic in place and to reinforce the sides of defective boxes. (Many of the box tabs were too short to snap into the slots after the plastic fronts were inserted, however, and the tape made the plastic insert very secure.) The finished lids were placed in rolling bins to be routed to the check-in line for reuse.

The check-in line occasionally had to be shut down for short periods of time because of the lack of empty bins, and at the peak of the check-in cycle there was such a shortage of bins that the ED box lids had to be loaded onto hand trucks.

When the bin containing the last of the ED questionnaires for a DO was bar-code read into the library, the bar-code station operator sent an "end" signal to the AICS. The computer program matched the receipts for the DO to a listing of what the DO should have shipped, and generated a "DO edit" list for review. Any discrepancies between the edit list and the list of shipped materials were referred to the Search Unit for resolution. Once the DO edit list was reviewed and accepted, the ICB was notified and ran a "DO accept" listing for the DO concerned. The DO accept caused the computer system to begin generating Staging and Transmittal Records (STR's) for the 100-percent processing workflows for the accepted DO.

Problems encountered—Some problems occurred during the receipt and check-in operation, including:

- 1. Late shipments, and highly compressed shipments, of questionnaires from the DO's.
- Inadequate staffing and space in the receipt and palletstorage area, and for moving rolling bins of material to operational units. (This problem was caused, in part, by the backup of material from the DO's late shipments.
- 3. Occasional lack of appropriate bar-code labels for DO shipments released to check-in or released by the ICB.
- 4. Slowdowns caused by referral of problem-ED forms to the Problem Resolution Unit.
- 5. Delayed training of some quality-assurance personnel, which slowed movement of ED materials.
- Difficulty in determining ED sequence when dealing with DO materials split between two or more shipments, once the DO edit listing was no longer available.
- 7. Inadequate staffing for the Plastic Inserter Unit.

Problem resolution—In a section devoted to problem resolution, the incoming bins from the check-in line were prominently labeled by DO, and the specific problems were identified, e.g., labels to be ordered, geographic problems, etc. When the bar-code labels were received, the repair clerks applied them to the appropriate ED boxes. The large number of labels being received sometimes caused the clerks to spend many hours cutting them apart and sorting them by DO. After about a month, the clerks were allowed to prepare handmade labels, which improved the flow of work out of the problem resolution section.

Every effort was made to preserve DO integrity, and the repair clerks tried to work on a specific DO and resolve the problem identified before moving on to the next one. During the peak of the check-in operation, this proved to be impossible, and the clerks had to work on many different DO's at the same time in order to complete bins of ED's. They routed completed bins to the bar-code station so they could be checked into the ED Questionnaire Library.

Geographic Problem Resolution

The Census Bureau published census data corresponding to geographic boundaries in effect on January 1, 1980. Because field operations had to start using the maps in 1979, they reflected January 1, 1978 or January 1, 1979 boundaries. The Bureau used these maps to define ED's for the census. To maintain flexibility in aggregating collected data into geographic tabulations, the Bureau designed ED's so that they did not extend into more than one statistical or political area. Boundary changes made after the maps had been prepared required the Bureau to split many ED's so that census data could be published according to the actual boundaries that existed January 1, 1980. There were about 37,000 such ED changes.

Geographic errors found during enumeration required correction of geographic codes and/or movement of questionnaires from one ED carton to another. Splitting ED's involved the DO's, the processing centers, the Data Preparation Division (DPD) in Jeffersonville, and Bureau headquarters. The Geography Branch in DPD provided boundary corrections. DO clerks transferred these to the enumerator maps, divided address registers so that they would coincide with the new ED's, changed the ED numbers on the questionnaires and the storage boxes (usually by adding an "alpha" suffix to the original ED number), and separated the questionnaires into new boxes. At closing, the DO's forwarded the boxes of questionnaires to the processing center along with a list of ED's for which changes had not been completed. The processing centers could not directly check in some of the boxes of questionnaires because the ED numbers did not agree with the numbers on the central list showing revised ED's. Processing center clerks diverted these boxes from the normal check-in flow. (See above.)

Where possible, clerks at the processing centers reviewed geographic problems and made the corrections in the problem-resolution section of the check-in operation. When cases could not be resolved by these clerks, or when a geographic problem arose later in the 100-percent processing operations, the relevant questionnaires were referred to the Geographic Problem Resolution Unit (GPRU), where geographic specialists checked ED boundaries and numbers, validity of block numbers, and other geographic problems, and made the necessary corrections. Some of the other problems referred to the geographic specialists were: ED's that had been improperly split by the district offices; ED splits and block-number changes that were required after the district office had closed; block-number changes that had been missed or done improperly by the

district offices; discrepancies between the block numbers shown on the address registers or questionnaires and those appearing in the master reference file (MRF); and ED maps missing from the address registers. The GPRU received a much heavier workload than anticipated. (A major reason for the increased workload was that some DO's (especially in Southern and Sunbelt States) closed without attempting to split ED's to reflect boundary changes, or to resolve some or all of the geographic problems that had been created or discovered during enumeration or the DO processing activity.)

CONTROL OF MATERIALS

Introduction

Each of the three processing centers received millions of questionnaires, thousands of address registers, and vast quantities of other records. To maintain control over these items, an automated inventory and control system (AICS) was created, using bar-code scanning as its major data-input system. (A manual control operation was used as a backup in case of failures in the automated system.) The AICS used bar-code scanning because of its ease in reading large quantities of data at a number of input control stations, each of which consisted of a cathode ray terminal (CRT), wand module (reader), modems, and, at three stations, a receive-only printer (ROP). There were 19 work stations at each processing office, each dedicated to the performance of a specific task in support of the overall system, such as check-in, library-in, and library-out.

Bar-code labels were printed both in house and by private contractors and subjected to quality-assurance (QA) procedures. The equipment necessary to the printing operation was a bar-code label printer, a CRT, and a keying station/console, operated in a remote-site terminal connected to a keying station under contract to the DPS. Correction labels were created through updates, additions, and deletions of ED (enumeration district) numbers in the data base. Additional labels were printed to replace missing, damaged, or unscannable labels. The bar-code label printer and associated hardware were manually operated.

Another control feature was the questionnaire and address register "libraries." The libraries provided a secure storage location for all materials; each ED had a location on the shelves in the libraries and always retained that location. This provided control (1) since only a small number of ED's had to be in the actual processing flow at one time, and (2) ED's could be found easily in their library storage location. The materials always were checked back into the library from any of the processing flows before being sent on to another. The libraries' operations were subjected to QA procedures, basically to minimize misfiling. (Three other types of libraries existed for the coordinated control of miscellaneous, reference, and film materials.)

Automated Inventory and Control System (AICS)

Introduction—The AICS tracked the movement of census materials from the time they arrived at the processing office through all phases of processing. As the material—ED boxes, address

registers, or film boxes—passed through the processing cycle, their location was monitored by having the information from the bar-code labels "scanned" and entered into the system.

In addition to capturing the workflow information, the barcode station operators edited (reviewed for accuracy) each box's markings to ensure that the item being processed was correct. Although the edit functions actually were performed by the centralized computer, certain manual steps were required of the operator, who interacted with the system by following instructions on the CRT screen.

The General Operations and ADP Support Branches were responsible for processing work units (each consisting of the contents of a rolling bin); the General Operations Branch organized and controlled the work units and resolved problems, while the ADP Support Branch operated the ADP equipment.

Equipment—The processing centers had several different commercial systems for providing remote-job-entry (RJE) support to their AICS operations. Production consisted of printing the staging and transmittal records (STR's), processing-status and quality-control reports, work-analysis edits, and inventory reports for ED's and AR's, and transmitting weekly quality-control data to Bureau headquarters. With normal service and maintenance, time loss was limited to solving problems such as nonrecoverable tape-read/write errors and sudden equipment failures or power interruptions.

General procedures—Bar-code clerks were responsible for scanning the contents of work units entering or leaving their stations. The work unit (WU) could contain ED questionnaire boxes, address register (AR) books, or film boxes, depending on the assigned station. Each bin was considered a WU, except those with film boxes, where each film box was a separate unit.

A Staging and Transmittal Record (STR) was created for each WU and placed in a pocket on each bin. In addition, three work-unit number cards were put in a triangular holder on the top of each bin (this permitted the work-unit numbers to be read from all sides of the bins) and a "bin transmittal card" showing "unit to," "unit from," etc., was attached to the side of each bin. Whenever a WU was read at a bar-code station, the operator entered the date and his/her initial in the "out" or "in" column of the STR, and returned the STR to the metal pocket on the side of the bin.

When the automated system became inoperative in the course of the processing operation, manual control (described below) kept the work moving. During normal operations, communication within the automated system was by means of a direct connection between the timesharing network concentrator and the work station. If the concentrator became inoperative, emergency operation provisions went into effect to bypass it via alternative "dial-up" communications linked directly to the time-sharing service's main computer complex. Six of the bar-code work stations were designated as critical processing

stations and were equipped with emergency "dial-up" capabilities. In the event of a failure, the operator would immediately notify his/her supervisor. The supervisor called the "Action Center," which would tell the supervisor the status of the system and the reason for the failure. (The reason for supervisory involvement was that various types of system failures could occur, and each situation required a specific procedure to be followed.)

AICS bar-code station operations—Automated data processing (ADP) supervisory personnel conducted training sessions for the clerks in this area with the aid of video films, training manuals, and on-the-job training with "hands-on" experience at the bar-code stations. Bar-code operators originally were hired to work mandatory 10-hour, 6-day work weeks when requested, but this requirement was later dropped because of the difficulty in locating personnel agreeing to work those hours.

Operator training began in July 1980, with 10 operators at each processing office involved in the first training program. Personnel from other ADP sections (Remote Job Entry (RJE), FOSDIC, and Camera) received cross-training in bar coding to assist in operating the bar-code stations. "Dry" runs were conducted continuously until the first DO shipment arrived for the check-in operation, and training sessions were conducted far enough in advance to allow familiarity with the task. ("Crash" training programs were conducted for additional personnel brought on board when required.) Originally, new operators were assigned to observe previously trained operations for 1 week, then assigned to their own stations. This observation period later was eliminated, and new personnel were assigned to a station for on-the-job experience.

At the start of the operation, the bar-code stations were set up in the following areas: ED Check-In, ED Library-In and -Out, Camera Prep, Camera-In, Camera-Out, Film Processing, FOS-DIC, Address Register (AR) Library-In and -Out, Diary-In and -Out, and Evaluations. Station locations were adjusted as the emphasis of the 100-percent operation went from check-in to library to camera flow to diary to coding to evaluations, and so on.

Bar-code operations started in the first week of August 1980. At various intervals, 10-hour shifts were run, overlapping into the next shift to cover absences, backlogs, training, and retraining sessions.

The staff had great difficulty in reading peeling, crushed, torn, or bent labels. Many boxes did not have the protective plastic fronts, and some ED boxes collapsed after being stacked too high. Sometimes ED boxes were returned to the library with STR's, bar code labels, and search data found in the boxes themselves instead of on the outside.

Address register (AR) bar-code labels were placed on the upper right portion of the back of each register. This positioning made it difficult to scan because all books were not placed flat on a table prior to scanning; some labels became loosened

in handling. Although the AR library had less material than the ED library, it had more problems with unscannable labels.

There were not many problems with unscannable labels for film boxes; handling was minimal and labels had to be handwritten in cases of multiple film boxes for a camera unit since the second box was never scanned. "Recycled" film boxes had their tops torn off when they came out of film processing so they could be identified easily. All film boxes were checked for label and film-box integrity.

Acoustic couplers, used when required to transmit information over telephone lines, frequently broke down. There were some problems in maintaining the bar-code stations, since each bar-code unit consisted of different pieces of equipment (e.g., a short-haul modem (modulator-demodulator), wand modules, regulators, cables, and a cathode ray tube (CRT)). Each piece of equipment was serviced by a different manufacturer and interacted with other pieces of equipment. Technicians were not present on site, so a call had to be made to the servicing agency; a delay of 1 or 2 days was not uncommon. Other equipment had to be mailed to the supplier for servicing.

The Special Functions station was not included in the original plans, but evolved to handle (1) ED's with zero population and housing units (ZPH's); (2) removing and/or replacing geography-hold status; (3) running ED and WU information; (4) locating missing boxes in check-in; and (5) printing STR's.

Bar-code operations revealed a number of problems related to particular stations' activities and organization. These were corrected by modifying procedures and adding staff and equipment.

Bar-code label printing—The ADP supervisor provided general operational and procedural instructions to the processing offices' staffs regarding the operation of the bar-code label printer, which was turned over to the experienced computer operators originally hired to operate other remote-job-entry (RJE) equipment. The operators had three primary duties: Operation and maintenance of the equipment; maintenance of the production, equipment-maintenance, and supply logs related to the operation; and the formatting of data into the form required for printing labels. No formal training was scheduled or required. Operation manuals were provided, and all personnel received on-the-job training.

The equipment required for this operation included a label printer, a keying station, and a console. This initial production of labels (those printed prior to the arrival of ED questionnaire boxes and AR's) began in the second week of March and continued until mid-May. Some labels also were printed by private contractors, due to time constraints and the limited capacity of the Bureau's printing equipment. The production of label reprints (of those labels rejected by quality control (QC)) began April 3, 1980, when the last QC operation (see p. 12) was set up. Printing the initial film box labels started May 20, 1980, and was completed June 19, 1980.

Test labels (for testing the AICS and bar-code labels in practical application prior to processing) were formatted in the

second week of June, and the printing started on June 20. Processing these labels was accomplished as the correction labels (those created through updates, additions, and deletions of ED's in the data base) were becoming available for formatting. Approximately 119,000 correction labels had been produced, covering almost all DO's, by August 13, 1980.

Printing replacements for missing, damaged, or unscannable labels began the week of June 9, and after August 13, all labels printed were replacements. The primary requesting unit was ED Check-In, for which about 35,000 labels were printed. Approximately 10 to 15 percent of all labels originally printed were replaced because they were unscannable.

After the printing of initial corrections and reprints, labels were matched with their appropriate printouts and sent to the Quality Control Unit. Replacement labels were keyed into the program, printed and attached to requests, and returned to the requesting unit.

The bar-code printer produced only six to seven labels per minute. Repairs or other required maintenance also caused delays; the response time for repair technicians was at least 3 work days and delivery of parts for the bar-code printer took between 2 and 4 weeks, even on priority orders.

Backup Manual Control Operation—In the event of an AIC system failure, a manual backup system was used. While in the manual mode, the bar-code operator verified the STR against the bin of work for all ED's listed. If no discrepancies were found, the STR was dated and the WU continued on to the next station. The operator annotated the WU numbers on the Recovery Log for the supervisor's "recovery" whenever the automated system was again operable.

Libraries

Introduction—The storage, maintenance, and cataloguing of materials at each of the processing offices was the responsibility of a system of libraries. Aside from the training library (discussed above in the section on training) and the film library (see p. 20 below), there were four documents libraries at each processing office—miscellaneous materials, ED questionnaire, address register, and reference materials. The functions of each of these units are described below.

Miscellaneous materials library—Census materials arrived by truck from the various district and other offices across the country. The trucks were unloaded and the materials sorted by type (ED boxes, address registers, and miscellaneous materials (e.g., used administrative forms, block header records, directories, and maps)) and placed on pallets or in rolling bins. The Receipt and Pallet Storage Unit stored the various materials until notified to send them to the appropriate processing unit. ED boxes and address registers were processed through their respective check-in units prior to being sent to the ED and

Address Register Libraries, while the remaining miscellany went to Miscellaneous Materials Check-In, where they were sorted and routed to the Miscellaneous Materials Library.

The materials were received at the library in rolling bins with form D-3328, Miscellaneous Materials Staging and Transmittal Request (STR), attached. Library staff verified the contents of each bin, checking it against the transmittal record; logged in the STR; sorted and counted the materials in each bin; and recorded the quantity received on the transmittal record. A copy of the record went to the keying staff where the information was keyed to the Master Inventory List (this list was arranged in district office (DO) order, and by form number within DO). The library control clerk then assigned shelf space for each bin of materials, and they were shelved in the library in the same sequence as the master list.

When a form D-3328 STR was received requesting materials from the library, the control clerk logged the request in the control log and, if the materials were available, they were brought to a temporary "holding area" before being sent to the requesting unit. (If the materials were not available, the STR was marked accordingly and returned to the requesting unit for resubmission later.) When materials were returned from a requesting unit, they were logged back into the library and reshelved.

ED questionnaire library—The questionnaire library at each processing office handled and stored approximately 30 million census questionnaires in 350,000 boxes. The questionnaires were sent to the library on a DO/ED basis as they completed ED check-in processing. Bar-code labels on the incoming boxes were scanned at a bar-code reader station to enter the identifying information into the computerized control system, the library control clerk logged in the arriving work unit, and the boxes then were shelved in ED order, by DO within State.

The ICB told the library when questionnaires were required for processing and produced an STR identifying the boxes of questionnaires to be formed into work units and sent for a specified processing flow. Each STR contained:

- 1. The "flow" (i.e., 100-percent microfilming, coding, sample diary, etc.) for which the materials were scheduled.
- 2. The work unit (WU) number.
- 3. The DO and ED numbers for all the boxes of questionnaires in the work unit.
- 4. The total number of boxes for the ED whose materials made up the work unit. (For all "100-percent flows," all boxes for an ED were included in the work unit being transmitted. For coding and sample "flows," only longform questionnaires were sent.)

Upon receipt of an STR, the control clerk logged the request and assigned library clerks to assemble the materials. They identified and removed from the shelves the boxes of questionnaires required, stacked them in a rolling bin (or bins) in ED-number sequence, placed the STR in the pocket on the bin, and sent the bin to the control clerk. The control clerk entered the date the work unit was "staged" (i.e., assembled and ready to be sent to a processing unit), and released the work unit to the bar-code clerk for scanning before sending it to the requesting unit.

When work units of questionnaires were returned from a processing flow, they were checked back in through the barcode reader station and given to the control clerk, who assigned the work unit (with STR attached) for reshelving. Once the boxes had been reshelved, the clerk involved returned the STR to the control clerk, who entered the date the work unit was reshelved on a control log and sent the STR back to the ICB.

Address register library—This library received conventional, master, and followup address registers (AR's) from the checkin unit. (These AR's contained the names, addresses, serial numbers, and types of questionnaires used for all housing units enumerated in the census, as well as the enumerator's maps.) The AR's were entered into the computerized control system by bar-code scan of their identifying labels, and then shelved in boxes in the library by DO and ED range. AR's identified during bar-code scanning as problem cases—e.g., extra AR's were present for a given work unit, or AR's belonging to a work unit were missing—were referred to the problem-resolution unit for corrective action.

Operational units requiring AR's for their work submitted requests through the ICB, which, as for questionnaires, generated STR's that identified the requesting unit and the AR's needed. When the library received an STR, the AR's requested were removed from their boxes, placed in rolling bins in sequential order, and sent to the bar-code reader station where their bar codes were scanned (to update the control system on their location) before they were sent to the requesting unit. Once the operational unit had finished with the AR's, they were returned to the library, their bar-codes were read again to check them back in, and the clerks returned them to the appropriate boxes.

Reference materials library—This library—one in each PO—housed all reference materials required for the coding operation, such as area-specific production coding guides, telephone books, city directories, maps, and so on. These materials, in conjunction with information from the questionnaires, were used to determine source information, verify addresses, etc. All reference materials were categorized and shelved according to type (place of work/migration (POW/Mig) by SMSA, generally in alphabetic (or numeric-code) order, and industry and occupation (I&O) by State).

The primary "customers" of the reference materials libraries were the POW/Mig, I&O, and general coding units. The coding unit (or other unit) supervisors requested materials as required and returned them to the library when they were no longer

needed. The libraries' control clerks maintained separate control logs for reference and training materials, and for administrative and other forms. The date requested materials were checked out of the library, the materials requested, and the date materials were returned, were entered in the logs. (Forms were not returned—they were "used up"—and the number requested was entered in the Forms Control Log.)

100-PERCENT PRETABULATION PROCESSING

Introduction

The 100-percent phase of the processing began when the PO's received the first complete shipments of questionnaires and other materials from the district offices (DO's) in August 1980, and continued until all the 100-percent data from all the short- and long-form (sample) questionnaires had been transmitted (see p. 23) to the Census Bureau's headquarters facility in Suitland, MD, and all diary review problems had been resolved. This task was accomplished on schedule in late December 1980, even though many district offices closed 2-1/2 to 3 months behind schedule. These data were used to reapportion seats in the U.S. House of Representatives. The 100percent data, also had to be available at the smallest geographic levels under the provisions of Federal law (P.L. 94-171), which required the Census Bureau to furnish population figures for appropriate political and statistical subdivision to each State government that wanted them for internal redistricting purposes. The Census Bureau also needed the 100-percent data (1) to provide the analytical tools for evaluating the accuracy of the data and (2) for use as the controls for weighting the sample responses to reflect the total population.

After being checked in, the census questionnaires were microfilmed and the developed film was sent to the FOSDIC (film optical sensing device input to computer) unit. The FOSDIC device "read" the 100-percent data and entered them on computer tape for transmission to the Bureau's computer facility at the Suitland headquarters for editing and tabulation.

100-Percent Microfilming Operations

General information—After the questionnaires were checked into the library, they were designated in the AICS (see p. 13) as available for microfilming on a scheduled basis. STR's were generated that grouped ED's into camera units (a "camera unit" was a group of ED's that could be filmed on a single 200 ft. roll of microfilm) and camera units into work units. (Rolls of film were spliced together in the darkrooms to produce the larger continuous rolls the film processing equipment accommodated.) A roll of film (a camera unit) could contain material for more than one ED, but ED's were not split on two different rolls of film. Following the STR's, the library clerks loaded bins that were routed from the library through camera preparation

and into the camera room. During the 100-percent microfilm flow, the 100-percent pages of the long forms and the short forms in their entirety were filmed. (Time factors prevented the microfilming of the 100-percent and the full sample questionnaires at the same time. The sample questionnaires required manual coding, and the Bureau could not do this and still have the 100-percent population counts to the President by January 1, 1981, as required by law. Thus, all of the 100-percent questionnaires and those pages of the sample questionnaires containing the 100-percent data were microfilmed first, to produce the apportionment counts. The sample questionnaires were microfilmed in their entirety after coding.)

After each roll of film was exposed, it was developed and checked for proper development quality and image density. If the film failed the quality control checks, the questionnaires for that roll were remicrofilmed.

While the film was being processed on FOSDIC, the microfilmed questionnaires were returned to the ED questionnaire libraries.

Camera preparation—Upon notification through AICS, ED work units (WU's-the contents of a rolling bin containing the questionnaire boxes) were routed to the Camera Preparation Unit (CPU). WU's arrived at the 100-percent camera preparation worksite either from the ED Questionnaire Library, containing ED's for 100-percent microfilming (new work), or from the 100-percent diary review for remicrofilming (i.e., ED's that failed the diary review operation and had to be remicrofilmed). The unit clerks prepared ED boxes of both long and short forms for the 100-percent microfilming operation, ensuring that (1) the WU contained the correct ED's in the proper box sequence, (2) a breaker sheet was in the first long-form box for each ED, and (3) that the ED boxes were grouped into camera units (CU's—i.e., the number of ED's sufficient to fill a complete roll of microfilm).

Upon arrival, the bar-code clerk checked each WU for validity. The clerk sent valid WU's to a control clerk, who logged the WU's in and routed them on to the breaker-sheet ID clerk, who located the correct breaker sheets (special sheets that listed each DO/ED, and a check digit) for the WU's and placed them on the bin. If an ED breaker sheet was missing, the ID clerk entered the notation "breaker missing" in the comments section on the STR and manually prepared breaker sheets for those ED's affected. An ED directory was used at the beginning of the operation as the source for the check digits needed in preparing the breaker sheets; this directory soon became outdated. Breaker sheet clerks frequently had to go to the Geography Unit to obtain the necessary information (the Geography Unit's directories were continually updated). If the information was not found there, a request was submitted through the ICB and the bin was held until a reply was received. The control clerk then assigned each WU to a camera preparation team.

The camera preparation team processed one ED at a time. A clerk removed the first long-form box for each ED, opened the

box, and verified the breaker sheet against the top questionnaire. The team verified the contents of each remaining box for
the ED by checking the top questionnaire against the bar-code
label. A CU breaker sheet was inserted into the first ED of each
camera unit to identify the ED's in each CU. The WU then went
to the Quality Control Unit, which carried out a sample verification of the contents of each bin. Any WU's with incorrect or
missing breaker sheets or other materials were returned to the
camera preparation team to have the problem(s) corrected.
After a WU was processed by the camera preparation team and
had passed the Quality Control Unit, it was logged out by the
control clerk and sent to the bar-code clerk for scanning. Upon
receipt of the "all clear" message, the WU was sent to the
microfilming staging area.

When 100-percent camera preparation began operations, the camera unit consisted mainly of microphotographers. Because there was great difficulty obtaining new employees at that time, the day camera teams were asked to prepare the WU's for the night camera teams to film. Later, enough clerks were hired to meet the needed levels of staffing. Camera preparation was a physically demanding job and some employees hired as camera clerks were unable to perform the physical tasks assigned; several laborers had to be added to the units to assist in handling the materials and equipment.

Camera operations—After the 100-percent CPU performed all required operations on a WU, it was routed to the 100-percent microfilm operation. During the 100-percent operations, all short forms and pages 2 and 3 of the long forms—the 100-percent data collected on both types of questionnaires-were photographed. In addition, ED's requiring remicrofilming because of diary edit failures were combined into WU's and processed through the remicrofilming operations in a manner identical to the 100-percent flow. Camera units failing film-processing quality control checks (recycles) were returned to the camera room and the associated ED's were refilmed.

WU's for microfilming were moved to the microfilming unit's receiving area where the control clerk, bar-code station operator, and assistant control clerk checked them into the unit and distributed them to individual camera teams on a first-in/first-out (FIFO) basis. These WU's were staged in specific rows designating either 100-percent, remicrofilming, or recycle.

When a camera (there were 22 cameras at the Laguna Niguel office, 20 at Jeffersonville, and 18 at New Orleans) was available for 100-percent microfilming, the jogging clerk received a WU and the associated film box labels (FBL's) and bar-code labels from the pre-camera unit staging area. The jogging clerk removed the ED boxes associated with each CU comprising the work unit and verified the breaker sheet. As each CU was readied to be filmed, the jogging clerk handed the camera operator an FBL and bar-code label, and an ED breaker sheet for each associated ED. The jogging clerk removed the questionnaires for each ED box, verified the contents against the

external label, and placed them in the jogging machine, which aligned the questionnaires for automatic feeding to the camera. The jogging clerk continued this process for each ED box in the CU. Upon completion of each ED box, and eventually the entire CU, the jogging clerk notified the camera operator and the unloading clerk.

The camera operator separately filmed the CU ID board (a board, filled out by the camera operator before filming a CU, showing the CU identification number, camera number, and date). Filming continued with the breaker sheet for the first ED, the first ED's questionnaires, the breaker sheet for the second ED, the second ED's questionnaires, and so on to the end of the CU. After the breaker sheets were photographed, the opened questionnaires were placed in a mechanical hopper attached to the automatic camera unit. The hopper fed the questionnaires onto a vacuum belt (the vacuum belt drew the pages of the questionnaires flat to ensure a proper image) that passed beneath the lens of the camera (called the camera head). Stroboscopic lighting illuminated each page for each exposure, achieving a "stop-action" effect that meant that the machines did not have to stop for each photograph. (This system was able to film up to 130 pages per minute per camera-a 50-percent increase in speed compared with 1970; in practice, the sustained rate for 1980 was closer to 80 pages per minute.) The unloading clerk removed questionnaires as they were filmed, placed them in their original box, and put the ED breaker sheet in the first long-form box for the ED. As each WU was completed, the unloading clerk returned it to the postcamera holding area. When the CU's filming was done, the camera operator annotated the film box label (FBL), placed a bar-code label on each film box, and the film in the box. As each WU was completed, the camera operator placed the film boxes on a tray located at the station.

The holding-area control clerk was responsible for the recording and staging of WU's, including delivering film boxes to film processing on a flow basis, keeping a record of filmed WU's and any recycles, and staging recycles for filming. Once informed by the outgoing bar-code clerk in the film-processing unit that a WU had passed film processing and was ready for release from the camera hold area, the holding-area control clerk had the WU returned to the ED Questionnaire Library.

Recycles-

The assistant control clerk periodically checked each camera station and removed the completed film boxes. The film boxes were then bar-code scanned and sent to the film processing unit, where the film was developed and subjected to certain quality checks, for such things as density and scratches. If the film for a CU failed the quality check, it became a recycle. The empty film box was returned to the camera room and the CU was refilmed. For recycles, the control clerk transcribed selected information from the original FBL to a new one, and placed the new label with the ED boxes for the camera unit to be refilmed.

The control clerk completed a transmittal at the end of each shift to cause new bar-code labels to be prepared for the recycles. The old film boxes were discarded.

Remicrofilming ---

ED's that failed diary review (see p. 21) were remicrofilmed after their questionnaires had been corrected. The remicrofilming process was the same as that used for the original microfilming, except that all control documents used were annotated "100-percent remic."

There were several reasons why an ED might be remicro-filmed, including: two questionnaires being filmed on the same frame, questionnaires misaligned on the film plane, one ED box in a multiple-box ED being left out, breaker-sheet failure, and so on. During the 100-percent remicrofilming operation, the WU's were sent directly to the camera preparation team, which verified that the ED breaker sheet was in the first long-form (sample) box and that the remaining boxes for the ED were present. If the breaker sheet could not be located, a new one was prepared and placed in the box. After the remicrofilm CU's had been processed by the camera preparation team and had passed the Quality Review Unit, the control clerk logged out the WU and sent it to the bar-code clerk. Upon completion of the bar-code scan and the receipt of the "all-clear" message, the WU was routed to the microfilming staging area.

In some cases there were WU's that contained zero population and housing (ZPH) ED's; these were identified by a special label, and all ED's associated with the ZPH work unit contained only breaker sheets for 100-percent microfilming. These WU's were handled in the same manner as other WU's.

The 100-percent microfilming operation was completed by December 27, 1980. Puerto Rico questionnaires were filmed between January 7 and February 27, 1981. (Those for the outlying areas were processed with the sample; see below.)

Camera maintenance—For the most part, the 1980 microfilm cameras worked well and proved very reliable. Mechanical problems encountered involved dust, the main drive system, coated light-sensing photo cells, and decreased strobe light output. In the main paper-moving drive system, including the feeder and stacker systems, there was the wear, tear, and misalignment expected with such complex equipment. Failures occurred in three areas, however, where they were not anticipated—the bearings, the main-drive vacuum belts, and the main-drive brake shaft—all of which required frequent repair or replacement. In the camera head itself, a major problem was shutter failure. Because the camera head was such a delicate mechanical device, a relatively high failure rate had been anticipated and the only difficulty was keeping enough in stock to replace the ones that failed. There was a related problem with the camera-controller circuit board.

Film quality was affected by such things as density and belt marks. Dirty camera lenses and light filters caused a decrease in image density. Normally, if a film had a low density, the processing area staff recommended an increase in strobe light output to compensate. This continued until no more adjustments could be made on the strobe unit, after which the lens and filter were cleaned and a "step test" determined the right strobe setting. Light reflecting from wear marks on the main bed belt and belt plate affected image quality. Physical damage to the questionnaires prior to their arrival at the CPU also caused problems. Tears, curled or folded edges, smudges, and so on sometimes were the result of handling by respondents, but also were caused by improper packing or handling at the DO's or the processing offices.

Film processing—The Film Processing Unit developed all the film produced by the microfilming, remicrofilming, and recycling ED questionnaires, and performed quality control checks on the developed film. Camera units of ED questionnaires were brought to the Film Processing Unit from the Microfilm Unit on trays and the contents were checked in at the bar-code station for verification by the AlCS by CU type, i.e., 100-percent, sample microfilm, remicrofilm, or recycles. After verification, each CU was sent to the film processor for development and quality checks on such matters as scratches, water spots, proper alignment of breaker sheets, and density readings.

Camera units that passed the quality checks were sent to the FOSDIC unit, while recycles went to the Microfilm Unit for refilming and then were returned to film processing for developing and quality checking. The old film was destroyed after quality-control clerks completed evaluations on the type of failure. During the course of the operation, the CU's were tested against FOSDIC specifications to maintain acceptable levels of developed film. Whether the film rolls passed or failed the quality control check, the bar-code labels on their boxes were scanned and verified against the AICS data base.

100-Percent FOSDIC Processing

General information—After CU's passed film processing, they were sent to the FOSDIC unit. In the first stage of the 100percent FOSDIC processing, the negative microfilm was "read" by the FOSDIC system, which detected filled-in circles (appearing as clear spots in specific places on each microfilm frame) on the breaker sheets and questionnaires that indicated responses or identification information. Each breaker sheet was scanned for specified information; if this was not complete, the breaker sheet and the entire ED were rejected. If the data were complete, the individual questionnaire pages then were scanned for responses. FOSDIC converted the coding marks into digital codes and electronically transmitted them in "real time" to the Census Bureau's central FOSDIC concentrator at Suitland, MD, where they were "logged" in and stored on computer tapes. Computers performed minor edits, imputed certain missing data, and generated the review diary (see below).

FOSDIC processing of the 100-percent data took place between August and December 1980. At each processing office, the FOSDIC Unit's responsibility was to transmit the data stored on the film into the computer system at Census Bureau headquarters using FOSDIC 80 computers. The computer technician at the processing center communicated with the concentrator (receiving) station in Suitland to connect the FOSDIC units for the electronic transmission of data. The FOSDIC operator loaded the computer programs from floppy disks into the processing office's computer following the instructions in the manual. The operator mounted the microfilm on the FOS-DIC; typed in the DO code, camera unit, and the number on the computer console; and began transmission of data. Upon a successful transmission of a film roll, it was bar-code checked out of the FOSDIC Unit and into the Film Library, where it was logged in and stored.

Staffing and training—ADP supervisory personnel conducted formal training sessions with the aid of video film and manuals and on-the-job training with the assistance of the computer technicians. The supervisor's training consisted of classroom sessions in census orientation, processing or FOSDIC operations, bar-code training, and training instruction. There was also on-the-job training. The operators were trained by the FOSDIC supervisor and the computer technicians, with handson training and manuals as reference guides. Because only experienced computer operators were hired, minimal training was necessary.

Each FOSDIC unit consisted of one supervisory computer operator and four computer operators—two per shift (the FOSDIC supervisor usually worked the "day" shift, while the general shift supervisor handled the "night" shift). Four FOSDIC machines were installed at each processing office in an environmentally controlled room and located so that an operator could run more than one machine at a time. Space was allocated for staging areas and necessary office equipment. During normal transmission hours, on-site computer technicians were available in an adjoining room.

FOSDIC operations—FOSDIC 100-percent film transmissions started in the first week of August 1980 at all three processing offices and the last roll of film was transmitted December 29, 1980. The FOSDIC program to transmit Puerto Rico microfilm was completed late in February 1981.

A microfilm input/output control area was designated within the FOSDIC unit to indicate which film boxes had been checked into the unit and were ready to be transmitted. When transmission was completed, the operator initialed the corner of the film box label that contained the DO and ED number. Transmitted-film boxes were then placed in the input/output control area designated for film boxes to be bar-coded out of the FOSDIC Unit.

There were various problems in the FOSDIC work area. Film labels were sometimes damaged or obscured, lacked information, or did not agree with bar-code labels. The supervisors

returned these film boxes to the Camera Room supervisor for correction. When the problem was resolved, the film boxes were returned directly to the FOSDIC Unit. Film was at times incorrectly wound on the reel, either backwards or upside down; it occasionally was damaged during transmission. The supervisor returned such rolls of film in their boxes to the film-processing supervisor for rewinding or refilming. There were some problems with modems, power sources, and transmission. Console messages on teletype indicated when any of these problems occurred. The computer technician would resolve the problems if they were local, or would notify Bureau headquarters if the responsibility lay there.

FOSDIC maintenance—Electronic technicians were hired and trained at the New Orleans processing office for both the New Orleans and Laguna Niguel sites (existing staff at Jeffersonville handled the FOSDIC equipment there). Once FOSDIC operations began, the technicians were responsible for equipment installation, new program implementation, preventive maintenance, repairs, and operator assistance and instruction.

Upon completion of training in mid-March, the technicians joined their respective processing office staffs and installed the equipment. Testing and resolving all the hardware problems continued into August, when processing of 100-percent microfilm began. The technicians were assigned to FOSDIC shifts as needed to provide maintenance and technical support.

Film Library

Once the rolls of microfilm had been processed by the FOSDIC Unit, they were sent to the Film Library for storage. Each roll in its own box with a bar-coded identification label attached was routed on a flow basis. The control clerk checked them through the library's bar-code scanning station, logged them in, and shelved them.

The microfilm storage shelves were divided into two groups, one for "100-percent" and the other for "sample" microfilm. The control clerks allocated shelf space by DO and State, and shelved the microfilm boxes in camera unit (CU) order within DO. The libraries received the first rolls of film for storage early in August 1980. Since these were in random DO/State order, there were continuous minor rearrangements as allotted space was taken up and DO/State groupings overlapped other groupings' space. The last film for the 50 States was stored in the libraries on December 29, 1980.

Only the FOSDIC Units, or the Film Duplication Unit at the Jeffersonville office, could request microfilm from the Film Libraries. Such requests were made through the ICB, which prepared Special Request STR's (form D-1331) and sent them to the libraries as required. The control clerk located the microfilm roll(s) requested and entered the CU number for the film, the date of request, and the requesting unit's acronym on the Film Library Request Log (form D-3281) before releasing the film to the bar-code station to be checked out. When the

microfilm was returned to the library, the control clerk updated the STR to show that the film had been returned and filed the STR in the "Special Requests Completed" folder, then returned the microfilm to its designated shelf area.

Diary Review

General information-After the 100-percent data were put on computer tape, and again after the sample data were entered, the counts for each ED had to pass a set of acceptance tests to make certain that data scanned by FOSDIC had not been lost or incorrectly recorded on tape, and that the potential errors or unusual entries did not exceed established tolerances for population, housing units, or various population and housing characteristics. Essentially, this involved comparing the initial computer counts with the 1980 field counts. The data that failed were summarized and printed out for each ED in a format called a "diary." The processing center clerks compared the data on the ED diary with the original questionnaires, address registers, and/or the microfilm to make certain that all forms were in good shape (not crumpled, too lightly marked, etc.) and properly identified, and that none had been missed during the filming process. For instance, the clerks might count the number of questionnaires in an ED box and then see if the same number had been photographed. When necessary, rejected ED's were remicrofilmed and sent through FOSDIC once again.

As it had in the 1970 census, the Bureau applied itemby-item tolerances as parameters in the FOSDIC program, so that any one (or all) could be changed during the census processing. Thus, once these changes had been made as a result of early review, the headquarters specialists concerned themselves more with such things as (1) systems errors that caused responses to be lost or incorrectly coded on the tape, and (2) a number of potential errors (for example, an excessive number of similar responses, such as "Other" to the race question). Inconsistency between the field and machine counts was a frequent reason for an ED to fail diary review. Geographic problems, such as an improper ED split, or ED's with missing or inappropriate block numbers, also caused ED failure in 100-percent diary review. (An ED had to be split when it was discovered that it was affected by boundary changes (see p. 13) or when more households and/or housing units were counted in an ED than could be processed in a work unit on the computer.) The cost for all diary review-100-percent and sample-was about \$4.5 million.

Diary review staffs were established at each of the processing offices, each consisting of an operations supervisor, a central control clerk, a processing analyst and one or more assistants, and a minimum of 8 to 10 clerical diary review units, each with its own supervisor, lead clerk, unit control clerk, and approximately 15 review clerks. Review clerks and the unit control clerks received formal training, using a videotape instruction program and exercises simulating the clerical tasks. The Diary Review Unit supervisors were responsible for providing staff on-the-job training.

The 100-percent diary review began slowly at all three offices in the first week of September, building quickly as the workload mushroomed with the concentrated delivery of materials from the DO's, and was completed late in December 1980.

Procedures—After the information on the questionnaires from an ED was transmitted to the Suitland computer facility via the FOSDIC operation, the data were computer-edited and the results returned to the processing offices in the form of "accept" and "review diary" listings (transmitted to the the remote-job-entry (RJE) terminal rooms at each office several times each day). The "accept" listings show the ED's that did not require diary review, and which remained in the ED questionnaire library available for other processing operations if necessary. The "diary review" listings identified the ED's for which the computer edits detected errors requiring investigation and/or correction. This information was also entered directly into the AICS via computer tape.

The central control clerks for the the Diary Review Unit at each office picked up the listings at the RJE rooms several times daily and distributed them to the review staff by DO (all the materials for a given DO usually were reviewed by a single unit). The Information Control Branch (ICB) at each office also received the STR's through the AICS, and routed the required ED questionnaire boxes and address registers in rolling bins from the various libraries to the diary-unit central control clerks. These clerks routed the bins through the diary review staff barcode reading station, where the bins' bar codes were read to update the AICS on the location of the materials, and then sent the bins to a holding area until one of the reviewing units requested the materials.

The control clerk in each reviewing unit matched diary listings assigned to that unit to the corresponding ED questionnaire boxes and address registers (AR's), placed the listing inside Box #1 of the ED boxes, the AR on top of the box, and placed the boxes on the reviewing unit's "ED's for assignment" shelves for easy access. The unit control clerk assigned ED materials to clerks for review as needed.

The review clerks were organized to perform the following specific types of operations:

Form counts—When the housing units and/or persons tabulated by computer for an ED were fewer than expected, a form count was required. The form count clerk(s) counted the number of questionnaires in the ED boxes and compared that number with the minimum number expected.

Block code review—Block code review included comparing block numbers for selected serial numbers on the diary listing to the block numbers in the AR. When the block numbers differed, the correct AR block number was entered in the diary. If enough block numbers on the listing were different from the numbers on the AR, the questionnaires involved were located and the questionnaire block numbers were edited.

Write-in review—This review dealt with the three 100 percent population questions requiring write-in responses (relationship, race, and year of birth). A diary review entry was made for each serial number (i.e., questionnaire) for which the computer edit program did not find a code for one or more of these items. The number of corrections required for each serial number was recorded in the diary.

Address register/diary (AR/diary) review—AR/diary review was carried out for ED's, previously reviewed by a processing analyst, that had significant differences between "edited" and "field" counts of population and housing units. The AR/diary clerk(s) compared the information from section IX (review operations) of the diary listing with information in the AR and recorded any discrepancies on the diary. Differences in block numbers, types of living quarters, and numbers of persons were noted for matching to serial numbers; serial numbers appearing only in the diary were noted, while those for addresses in the AR were written on the diary with their block numbers, type of living quarters, and population numbers.

Worksheet review—Worksheet review was done only for ED's routed to an AR/diary clerk. The worksheet clerk tallied the differences reported by the AR/diary clerk and summarized the results on a worksheet. Significant differences between the AR and the diary required a questionnaire edit and remicrofilming of the ED.

Questionnaire edit—This edit was done only for ED's subjected to AR/diary review. The edit clerk checked questionnaires with serial numbers that the diary listings indicated required review. (These were identified by handwritten entries and/or by block number, serial number, or population-number suffixes on the diary.) The diary information on these serial numbers was compared with the corresponding questionnaires, and the clerks used editing tables to make corrections to the questionnaires, diary, and/or the AR.

Group quarters²—ED's with significant differences between the edited and field group-quarters population counts were reviewed by the group quarters clerk, who compared the information in the diary listing with that on the AR and entered any discrepancy on the diary. Significant discrepancies required a questionnaire edit and remicrofilming.

One or more clerks in each diary review unit were assigned to each of these functions. Each clerk received review assignments from the unit control clerk. When each job was completed, the clerk indicated further action (if any) to be taken in section VII on page 1 of the diary review listing. If changes were not needed, or were so slight as would result in a negligible change in the data for the ED, the clerk circled the word "Goldplate" in section X on page 1 of the diary review listing and initialed the page. "Goldplate" ED's were "accepted" and did not require further review; the "Goldplate override" was keyed and the diary flag removed so that the ED was available for further processing.

Materials for diary review might have to be reviewed several times by different clerks before passing out of the unit entirely. For example, an ED that underwent an analyst's review (see below) might have to undergo an AR/diary review and a questionnaire edit as well.

When the individual clerks completed their parts of the review, they placed the ED boxes and/or AR's on the "materials check-out" shelves in the unit for staging to the appropriate destination—either to another review clerk or outside the review unit. Remicrofilming and "Goldplate" materials were collected in rolling bins, by DO, for staging, while materials requiring unusual corrections or changes were sent for procedures-analyst review.

Referrals—Clerical review procedures did not cover all possible problems and specified cases were sent to the diary review unit supervisor or to the procedures analysts for resolution. Supervisor referrals included ED's with damaged or missing questionnaires, serial numbers not found or cancelled in the AR's, questionnaire edit and group-quarters referrals, Spanish-language questionnaires and other forms, geographic problems, and control-clerk referrals. The supervisor reviewed the referral cases and determined what actions were needed.

The procedures analyst had to deal with three major types of referrals: (1) Group-quarters problems, (2) ED's with clerical errors, and (3) "other" referrals. Groups-quarters referrals to the analysts were those not covered by the supervisors' instructions. Normally, the attached "shuttle card" had the information required to reconcile discrepancies in the data, but if not, the analyst had to decide how to resolve the problem based on his or her own experience and general census procedures. Materials with clerical errors required a review of the diary listing and associated questionnaires, and correction as necessary. "Other" referrals usually were from diary review unit supervisors seeking advice or assistance in resolving particular referral problems of their own.

Quality review—A quality-review program was used in the diary review units to detect and correct problems in clerical performance, procedures, or the training program. Each unit's lead clerk was responsible for daily quality checks on the performance of each review clerk. The lead clerk, sometimes assisted by the procedures analyst and his/her assistants, checked the first few ED's processed by each review clerk in each functional position. Review clerks completing two successive acceptable ED's with the same review action (i.e., AR/diary,

²Group quarters included institutions such as hospitals and jails, and non-institutional group quarters such as missions, dormitories, and other living quarters shared by nine or more persons unrelated to the person listed in column 1 on the "first form" questionnaire.

block code, worksheet, etc.) were considered qualified for that review action. Thereafter, one example of each qualified clerk's work was checked daily, unless a clerk was assigned to a different review action, in which case he or she had to be requalified. A clerk having two or more diaries rejected for quality reasons out of 10 consecutive diaries was disqualified and assigned for retraining in problem areas, as well as "warned." Any clerk receiving two disqualifications in a 30-day period was reassigned to write-in or form-count clerical review (considered less demanding than the other positions) or, if already in one of those positions, was removed from the diary review staff.

Diary review closeout—The Census Bureau is required by law to complete the final census population counts needed for congressional reapportionment, and to transmit them to the President within 9 months after the census date. To meet this requirement, the agency developed a schedule of proposed closeout dates for processing the data from each State, so that at a specified point in processing each State's data, the 100-percent data corrections could be halted.

The first States closed out through diary review were Vermont (October 28) at Jeffersonville, Delaware (October 29) at New Orleans, and Hawaii (November 24) at Laguna Niguel.

The 100-percent data processing at the processing offices originally was scheduled to close down on December 15, but by late October it was obvious that the late closing of the field offices and the attendant delay in processing could change this date. The Decennial Processing Staff (DPS) began transferring personnel within the offices from coding and other activities to concentrate on the 100-percent operation, Actions of persons and institutions outside the control of the Census Bureau caused some of the delay, notably in completing local review of some ED's and resolution of ED's in several major cities involved in litigation concerning the enumeration. Within the diary review operation, three principal problems were encountered: (1) "Goldplate" ED's that were not promptly sent out of diary review; (2) ED's scheduled for remicrofilming that the AICS showed still were in diary review; and (3) an undetermined number of ED's sent for remicrofilming that were returned to the ED questionnaire libraries instead. These difficulties were primarily the result of work backlogs caused by delayed delivery of materials and higher-than-anticipated rates of corrections required. The AICS proved extremely useful in locating materials that had been misrouted or delayed in movement, and the processing offices were able to address the problems relatively easily. No major changes in procedures were required. The last States to complete diary review at the New Orleans office were Florida, Louisiana, Texas, and Virginia-all on December 20. Jeffersonville completed diary review of all the States in its processing area except New York by December 13. New York was delayed because of the reenumeration of the Bedford-Stuyvesant section of Brooklyn (see ch. 5), but diary review of the State was not completed until December 24. The Laguna Niguel diary-review operation closed out the last three States handled there—California, Missouri, and Oklahoma —on December 28.

SAMPLE PRETABULATION PROCESSING

General Information

This section describes the processing of the sample data derived from about 15 million long-form (sample) household questionnaires. The 100-percent and sample responses were processed in the same manner, except that handwritten answers to many sample questions had to be converted into machine-readable codes. Because detailed data classifications were needed for a number of the sample items, such as income, occupation, industry, place of work, language, and ancestry, respondents had been asked to write in exact answers instead of marking a choice on a list of possible responses. Coding these handwritten responses was the most time-consuming and expensive procedure at the processing centers, involving about 3,000 clerks and \$27.2 million between January and October 1981.

Three separate sections of clerks did the coding. One section worked solely on the place-of-work (POW), travel-time-to-work, and migration questions (known collectively as "POW/Mig"); a second dealt with the industry and occupation (I&O) answers; and a third section coded all other general items, such as place of birth, language, ancestry, income, homeowner costs, and the like. Using various reference materials (e.g., coding guides and telephone and ZIP Code directories), the clerks determined the codes for the handwritten responses and then filled in the appropriate circles in the designated code spaces on the long-form questionnaires. (See ch. 12 for illustrations of the individual items and their coding boxes.)

Because any large clerical operation produces errors, a quality control operation checked the clerks' work. Specialists encoded selected questionnaire samples for verification purposes. At times, this process involved comparing separate codings of the same data and, when the error level was deemed unacceptable, correcting the errors. Depending on the workflow, backlogs, and problems encountered, a group of questionnaires could stay in the coding sections for 6 weeks or more.

Once the sample questionnaires for a group of enumeration districts (ED's) were coded, the sample data were transferred onto magnetic computer tape via the same two-step system that was used for the 100-percent data: The questionnaires were grouped by ED and microfilmed by specialized high-speed cameras. Then FOSDIC (film optical sensing device for input to computers) scanned the microfilm and converted the coding marks into digital codes. (See p. 19.) The four FOSDIC machines in each processing center transmitted the sample data to Suitland, MD, between July and December 1981.

Because of budget constraints, the Bureau decided to slow down the sample coding operations beginning in early March 1981 and stretch out the work into the next fiscal year (beginning October 1, 1981) by reducing the staff. For the same reason, it was decided to code the POW/Mig responses on only half of the sample questionnaires, thus reducing the sampling rate for three items (but increasing their sampling error by 40 percent). When the budget situation improved in June 1981 with the appropriation of supplemental funds, coding was accelerated by restoring staff, but the POW/Mig coding reduction was retained.

Clerical Coding of Handwritten Responses

Coding operations began with the receipt of ED's from the library. Whenever possible, POW/Mig and I&O coding were scheduled independently, but the materials for both were grouped at the State or SMSA (standard metropolitan statistical area) level rather than at the ED levels needed for general coding.

Each sample questionnaire requested information on up to 7 persons, with a possible 39 questions for each. Of these, about 15 called for written-in answers. POW/Mig coding involved 2 of them; I&O assigned codes to 4 from 503 possible codes; and general coding was responsible for the remaining handwritten responses, including the 7 possible for the housing unit. In some instances, the coder had to make simple calculations to arrive at the proper code, such as converting a monthly income figure into an annual total. The general coders were also responsible for affixing a piece of black tape in a designated place on the last page of data for each sample questionnaire. (This "shadow bar" tape, detected by an electric photosensor on the automatic camera unit, triggered the documenttransport mechanism to move the completed questionnaire off the camera bed and position a new questionnaire for filming.) A special clerical unit separated and marked a FOSDIC circle on the sample questionnaires not coded for POW/Mig so the FOSDIC operation would accept them.

Constraints on the sample coding operations—There were several delays in the coding operations. Coding was scheduled to begin with a gradual buildup of the staff in October and November 1980, with full production to start by January 1981 and be completed in June or July 1981. However, staffing did not begin until January 1981, and the full complement was not operational until March. As noted above, budgetary problems then forced reduction of the coding staff by approximately one-half, thereby extending the schedule even further. By the time the staff had been restored to full strength in June, the target date for completion had slipped 3 to 4 months. Part of the reason for the delay was the long time necessary to train a production coder and have that person reach maximum efficiency. The trained coders released in March often were not available for subsequent rehiring, so that new ones had to start at the very beginning of the training and coding process.

The overall extension of the coding schedule inevitably meant delaying publication of the sample data. To meet user needs, the Bureau decided to process and tabulate a 1-percent national sample of population and housing data ahead of all the rest. This "Early National Sample" (see ch. 8) had high priority and caused minimal disruption of the normal coding operations as selected ED's and reference materials moved through the system.

Quality Control

"Dependent" and "independent" verification procedures in conjunction with the sample coding operations maintained quality control of the production coders' work. In "dependent" verification, the quality-control coder, called a "verifier," saw the production coder's work and potentially could be influenced by it; in "independent" coding, two coders, who did not see each other's work, coded the same questionnaires, and their work was compared.

The general coding operation used dependent verification: A verifier first selected a sample of questionnaires that a production coder had processed. The verifier then marked on the quality-control record whether he/she agreed with the production code; if not, the correct code(s) was entered on the form and the one(s) on the questionnaire was changed.

The POW/Mig and I&O coding operations used independent verification. Here, a "precoder" coded the responses on a selected number of questionnaires onto a special form. Next, a production coder assigned the codes directly onto the questionnaires. Then a matcher retrieved the forms and the relevant questionnaires and transcribed the production coder's decisions onto the precoder's form. If they matched, no further action was necessary; if they did not, a postcoder dependently coded the response and entered that code on the precoder's form. The three codes were compared; the two that matched were considered correct and the odd code was charged as the error. (If none of the codes matched, no error was charged.)

An acceptable error rate for POW/Mig coding was 5.25 percent on an item basis. If a coder's work units exceeded this rate, he/she was warned. If the rate was exceeded on three out of any five consecutive work units, the coder was retrained. If retraining was required a second time within 20 work units, the coder was to be removed from the operation.

Production standards, set 5 to 6 weeks after coding began, ranged from 47 sample persons per hour for large SMSA's to 100 sample persons per hour for non-SMSA areas. Coders were timed again after these standards were established, but no significant differences were detected.

Microfilming of Coded Questionnaires

Five months elapsed between the end of 100-percent microfilming and the beginning of sample microfilming. By then, roughly 90 percent of the camera operators and many of the supervisory personnel had either left the census or were engaged in other work, so that a practically new work force had to be organized. Recruiting two operators for each camera proved difficult, particularly since the microphotographers' positions had been changed from wage grade to clerical, effectively cutting pay in half. Those operators that were hired, many of whom now were general clerks, received 2 days' on-the-job training from experienced operators, using practice questionnaires. Fast learners were teamed up with trainees that required additional instruction. Arrangements were made with the ED library concerning the number of work units it would release each day, and logs were set up to control movement of the color- and bar-coded work units as they were processed and to record progress through preparation and filming. Production was low at first, but with relatively few recycles, it rose weekly. As more staff were assigned, the former trainees became the trainers.

After coding and quality control, the sample questionnaires, including the pages already photographed in the 100-percent microfilming process, were microfilmed in their entirety. As noted above, sample microfilming followed the same basic procedures as the 100-percent operation, except that an automatic page-turner was used. When a sensor detected a piece of black tape on the last completed page of a questionnaire, the document-transport mechanism ejected the questionnaire and brought a new one into position under the camera lens.

Several changes were instituted in the camera-preparation operation for the sample. The two-person teams were broken up, so that each clerk worked separately. A special clerical team prepared the density sheets (identifying the district office and film roll numbers) for the camera operators, so the operators no longer had to process these sheets in their darkened work areas and could concentrate on camera production.

The first test run of sample microfilming began in mid-May with all the ED's for Montana, to see what personnel needs might be for the operation. The Montana work units, however, were in almost perfect readiness, with the forms in the ED boxes evenly placed, almost every sheet in proper position for filming, and nearly all the breaker sheets in their assigned boxes. The second State, Wyoming, proved to be more typical, with torn and partial forms, too-heavy taping, breaker sheets missing, and the like. Correction of these conditions sometimes introduced other problems, such as serrated edges or cuts or mars on the documents, that caused loss of production. Camera operators helped the preparation clerks as needed until there were sufficient work units for the cameras. Unit supervisors maintained daily diaries detailing the work done.

By mid-July, roughly half of the cameras were in full-time operation and most of the Early National Sample had been processed. The ED library began to increase substantially the number of work units released for camera preparation, leading to the need for more area for storage and longer work hours—10 hours a day for both camera preparation and processing. Work continued at an accelerated pace throughout the summer of

1981. The first work units for the outlying areas (which did not have a sample but were processed at this stage) arrived in the camera area in early September. They, together with the Puerto Rico sample, were virtually all microfilmed by the end of November 1981, when the nonsupervisory camera personnel were released.

Various problems were encountered that slowed the sample microfilming operation. Shadow bars (pieces of black tape) on the questionnaires sometimes were absent or misplaced and the questionnaires themselves had been packed in a haphazard manner. Breaker sheets occasionally were absent, incorrect, or mixed in with the accompanying questionnaires. Boxes of ED's that coding units had referred to analysts for problem resolution sometimes found their way into camera preparation instead. Voltage fluctuations required replacement of fuses and relays.

Film processing, as described on p. 19 above, followed.

Film Library

File cabinets were relabeled to accommodate sample microfilm, which required 40 percent less storage space than had been necessary for the 100-percent film. A remote printer, used with a bar-code station, replaced three control clerks and proved to be a faster and more effective method of controlling and recording data. Otherwise, responsibilities were the same as for the 100-percent film.

The film libraries received the first rolls of sample film for storage in mid-May 1981 and the last by mid-December. The film library supervisor filed them on a continuous basis and encountered no major problems or backlogs. On November 10, 1981, operations personnel began film quality control (QC) and shipping. Three separate shipments, on November 17, December 1, and December 17, were made from Laguna Niguel and New Orleans to Jeffersonville. There was a delay in the receipt of the computer listings from headquarters; otherwise, there were no problems encountered in meeting deadlines.

Sample FOSDIC Processing Operations

FOSDIC sample-data processing consisted of transcribing the encoded responses to a magnetic tape, with review and clerical repair taking place after computer editing. All other FOSDIC operations for the sample data were the same as those for the 100-percent data (see p. 19).

One supervisory computer operator and a computer usually constituted the entire operation. Bureau headquarters supplied the necessary programs, which were stored on "floppy" disks and used when transmitting the data for a particular area—a State or (in the case of Laguna Niguel) Puerto Rico or an outlying area. No other special preparation was necessary.

Sample-data processing on FOSDIC began in mid-May 1981. Transmissions were intermittent through July and into August but intensified thereafter, reaching their peak in October. FOSDIC operations were completed and shut down during the second week of December 1981.

Control of the boxes of sample film was handled in the same manner as for the 100-percent film, and film processing differed only in the programs used. The total volume handled during sample FOSDIC processing equalled 57 percent of the 100-percent processing workload.

Sample Diary Review

Diary review here was performed in much the same manner as for the 100-percent data (see p. 21 above), except that it involved far more detail given the number of data items being tallied and cross-checked. Sample diary review did not deal with population and housing counts, but rather focused on the presence or absence of written-in entries, and on cases where the FOSDIC system had not detected a code marking—either because of failure in calibration or faulty film processing. When the ED-by-ED review revealed an excessive number of errors, clerks had to correct them manually. Diary review for the sample began in June 1981 and ended in December 1981.

OTHER PROCESSING CENTER OPERATIONS

Automated Inventory and Control System (AICS) Operations

Label printing—Film-box label keying and printing for the sample began in the second week of April 1981; any labels that the quality-control operation rejected after scanning them for validity were recycled through the printing system until accepted. After this, production turned to the ongoing function of replacing labels for address registers and ED and film boxes. An average of 50 replacement labels a day had to be generated. All labels carried district office numbers and completion dates, and production logs were maintained for them. Replacement labels were sent to their requesting units.

The equipment problems that had hampered the 100-percent label-printing operations (see p. 12) had been resolved, and the label printers were not subjected to the same workloads and extended periods of operation necessary before. Lower frequency of problems and loss of time while waiting for service and parts appeared to have a direct relationship to the lower production requirements. The labels were printed far enough in advance of the sample camera operation that the limited printer speed—six to seven labels per minute—and temporary losses of system accessibility were not significant handicaps, but some reprinting was necessary because of computer programming problems, insufficient equipment cleaning and maintenance, and the use of overage print ribbons.

Keying Diary Review and Other Programs

Each processing office had a special area for keying operations. It normally housed 30 or so keying stations, 2 tape drives, 2 line printers for receiving transmissions, 2 disk drives, a small tape library, and a fairly large document-staging area. Keying personnel, including control clerks, were trained in a variety of ways. New keyers were hired as needed, and when a night shift had to be added, it usually was composed of ADP (automatic data processing) personnel specially trained for particular operations.

Diary review—"Diaries" were computer-generated listings of instances where the data for ED's and on individual question-naires within them did not meet established tolerances for the collected data. In "diary review" (see p. 21), clerks examined the listings, compared them with the source documents, and made corrections or repairs as appropriate. This could include remicrofilming all the questionnaires in a given ED and repeating the subsequent processing. Actions that did not require processing were reported through keying, such as the following:

Goldplate—The ED was to be accepted on the current cycle, because no significant improvement would result from processing it again.

Revert—Data on the diary had been edited in a cycle prior to the latest one because the ED had gotten into the wrong processing flow (e.g., to the camera instead of the library).

ZPH—In 100-percent processing, this indicated an ED with zero population and housing; in sample processing, it meant that no sample data had been collected.

Keying for 100-percent diary review took place between September and December 1980; for sample diary review, it began in June 1981 and ended in December 1981. The diary-review data were keyed to tape and transmitted to headquarters daily.

Post-enumeration survey—The 1980 PES program (see ch. 9) employed sample-survey methods to measure directly the components of census error for a sample of persons, and by this means to estimate the net error for the census. It involved (1) matching information collected in the periodic Current Population Survey (CPS) taken in April and August 1980 with the census to see whether the persons in the CPS households were also enumerated in the census, and (2) interviews of approximately 100,000 enumerated households chosen by sampling the census records for about 10,000 ED's and generating a control list by computer. The clerical operations were carried out in the three processing centers beginning in September 1980 and required access to the prelist, master, and followup address registers, and to the census household questionnaires themselves. Clerks located the CPS and PES sample addresses in the census registers, identified their census serial numbers and form types, located the appropriate questionnaires, prepared interview forms and sent them to the appropriate regional offices, received the completed forms, and

matched the sets of CPS, PES interview, and census records. The match results were edited and batched in groups of 99 or fewer documents and forwarded for keying.

The data on the source documents were keyed twice, independently, and transmitted through the Bureau's communications system to headquarters, where a computer compared the two sets of data and generated a "correction listing" showing the instances where the two keyings were not identical. Keying instructions changed frequently as analysis revealed faults in the procedures. Clerical resolution, keying, transmission, and computer editing continued in cycles until the file reflected actual differences between the census and survey data. The bulk of the keying was completed by February 1981, and most of the staff were assigned to other work by May 1981.

Other programs—To provide management with quality and production information, data from both precode and code records were keyed to tape and verified daily, and transmitted every Friday between December 1980 and October 1981. Often, keying was the fastest way to ask for and provide information on address registers and ED boxes. Initial requests for that sort of keying began in November 1980 and continued to mid-May 1981.

Count Review

After the apportionment totals for each State had been delivered to the President at the end of December 1980 and some of the preliminary and advance census reports had been issued (see ch. 8), processing center clerks reviewed the 100-percent ED population and housing counts within 40 of the 50 States by comparing them once more with the preliminary figures the district offices had prepared. Significant differences were investigated and reconciled, usually by correcting discrepancies in geographic-area designations, allocation of population and housing units to ED's, and the like. These corrections, resulting in changes to within State totals but usually not in the State totals themselves, were carried into subsequent tabulations.

CLOSING THE PROCESSING OFFICES

All of the address registers and other records were shipped to Jeffersonville in January and February 1982, and furniture, equipment, and property at Laguna Niguel and New Orleans were returned to the Commerce Department or to lessors through the Government Services Administration. Shortly before the processing centers were closed, most of the questionnaires and other confidential materials were destroyed under security conditions and reduced to pulp for recycling. A small percentage of the documents were retained in Jeffersonville for sampling and evaluation purposes (see ch. 9), but because of extensive litigation over the census (see ch. 10), questionnaires

from 136 district offices had to be kept in Jeffersonville during the ensuing decade. (That facility stored over 600,000 address registers from the 1970 and 1980 censuses and the cameras, film-processing machinery, and FOSDIC equipment as part of its normal function.) The sample microfilm was stored in the Jeffersonville facility's film library, but the 100-percent microfilm (including the 100-percent pages on the sample questionnaires) was duplicated. As required by statute (Title 44, United States Code), the original set was deposited in the National Archives, where by interagency agreement it would be closed to the public for 72 years (i.e., until 2052). A duplicate set remained in Bureau custody for use in the "age search" operation at its Pittsburg, KS facility. Under the confidentiality provisions of the census law (Title 13, United States Code; see app. 1A in ch. 1), members of the public could apply there for official transcripts of their own census records for the period 1900 to 1980, or for other individuals' records for the same period with appropriate authorizations from the named persons. (The National Archives' record holdings from 1790 to 1910 were open to the public.)

The New Orleans and Laguna Niguel processing offices and the decennial processing operation in Jeffersonville (except for the storage of records there as noted above) were closed officially on February 26, 1982.

HEADQUARTERS PROCESSING OPERATIONS

General Procedures

At Bureau headquarters, the census data received from the processing centers were placed on computer tape, "run" using a series of editing programs that (1) ensured that the data recorded from a questionnaire reflected actual responses and not just stray marks, (2) eliminated inconsistent data, and (3) provided data missing on the questionnaires. The data were further checked for accuracy, completeness, and consistency through professional review, and were eventually tabulated and formatted for publication. (The publication process is described in ch. 8.) Although data began arriving at headquarters in August 1980, tabulation beyond simple counts did not start until January 1981, when all of the 100-percent data were completely captured and accepted. Tabulations for a given State could not begin until all of its data had been processed up to that point.

The computers processed the data for each person and housing unit through statistical editing and (for sample data) weighting routines on the basis of very detailed specification programs. The intent here was to make the published statistics more accurate in their description of the population and its housing, and more useful than if "not reported" categories were added to each tabulation instead. After editing, each sample person and housing unit was assigned a weight to produce estimates of the figures that would have resulted had

all of the Nation's households responded to the sample questions. (See ch. 7.) When these processes were completed, the edited data about all persons and housing units, with their geographic identifications and weights, were stored on computer tape.

Detail-File Preparation

The detail file—a group of individual person and housingunit data records—was created during the editing operations, and was subsequently used to tabulate the summary tape files (STF's; see ch. 8). The detail file was a final record of the population and housing-unit data merged with the final geographic boundaries. As it was created, geographic files were produced, the data were tabulated and cross-checked against the geography, and editing occurred. (Final maps were drawn, but in an operation separate from processing.) The basic purpose of the detail file was to ensure (1) that all the data had been captured and checked for accuracy, and (2) recognition of all appropriate ED's and the correct block numbers within them. The two basic elements in creating the detail file were (1) transferring census information collected in the field through the FOSDIC operation (see above) and checking the results, and (2) tabulating the figures and editing and weighting the results. When tabulating the detail file, it was discovered that some ED's had not been split or had been split incorrectly. This meant that these ED's had to be split either in the field, the district office, or the processing center. If such situations had not been corrected early enough, they were changed in formulating the final, edited ED files by computer.

Data for the detail file were stored in two phases—100-percent and sample—on a flow basis as they were transmitted from the processing centers. The 100-percent processing took about 4 months to complete, as many geographic boundary reviews and statistical cross-checks ("editals" and "analyzers") were necessary to correct all the data before further tabulations were made.

Creation of the sample-data detail file generally went more smoothly than the 100-percent file, because the geographic problems associated with the latter had been resolved. Sample processing was complicated, however, by the additional editing, allocation, and weighting procedures that had to be completed before the data could be released for tabulation. There were several important weighting factors (see ch. 7); one was sample weight, which was the initial one assigned to all sample data to equate them with 100-percent levels. In the case of the subsampled data from the place-of-work and migration responses, only half of which had been coded for some areas (see p. 24 above), a compensating, supplemental weight had to be assigned.

Edit tallies were compiled from the detail file and reviewed by subject-matter analysts; tabulation tallies were generated after the edit tallies had been cleared. Edit tallies were in a table format designed for internal audit rather than external use; they also were at the State level, whereas tables went to the census-tract level. Only tabulation tables and analyzer programs—both produced at the same time—were used in the review process. Edit tallies were only used to check initial detail-file data. Generally, errors found in the edit tally review were the result of incomplete processing (such as missing ED's), but they also could indicate problems in the edit specifications.

Once the 100-percent detail file was completed, computer processing began for STF's 1 and 2; STF's 3, 4, and 5 followed after the sample detail file had been tabulated. Subsequent production of statistics was not as difficult as getting the "raw" numbers into the processing system correctly, as many computer programs could not be tested and adjusted until large amounts of data had been collected and processed. Problems often did not surface until then.

Computer Edits

Despite the efforts in the field to complete all the questionnaires, there were still some housing units or persons for which certain characteristics could not be determined. In addition to the edits, discussed above, that involved checks to make sure that the information recorded for a questionnaire reflected actual responses and not just stray marks on a page, there was a series of computer edits that accounted for inconsistent or missing data. For example, a householder could not move into a unit in a year before the structure was built or before the householder was born, so the computer changed the year moved in to an appropriate interval. An answer to replace a blank for the same question would be supplied ("allocated") from a preceding unit of similar age and tenure. (See ch. 12 for a discussion of allocation and its application to specific questions.)

Another technique, "substitution," was sometimes used. This occurred when a person or housing unit was known to be present, but no characteristics had been recorded because no interview had been possible or there had been a mechanical failure in the processing system, such as illegible microfilming of a questionnaire page. In such cases, the computer selected data from a previously processed housing unit as a substitute, and it duplicated a full set of 100-percent characteristics for the unit and each occupant.

There were specific tolerances for the number of computer allocations and substitutions permitted for any given geographic area. If the number exceeded the tolerances, this was noted on the diary, and resolved in diary review as described earlier in this chapter. Certain printed reports and most summary tape files (STF's) included tables showing the amount of allocation and substitution for particular items.

Sample Weighting

Following computer editing, the sample data were subjected to a procedure that assigned a weight to each person and housing unit. For areas sampled at a 1-in-2 rate, the sample

weight was close to 2; in areas sampled at a 1-in-6 rate, the weight averaged about 6. Thus, to obtain tabulations for any characteristic for a particular geographic area (e.g., the number of persons in Elm County in a particular income category), the weights for the sample persons and/or housing units with the characteristics of interest were simply summed. Further, the weights were assigned in such a fashion that for most large geographic areas, the 100-percent census counts and the sample tabulations for total population and total housing units were very close. The weights or multipliers, when applied in stages to the sample data for any given area, produced figures that matched or were very close to the complete-count figures for total population, race, sex, age, family size, and certain housing characteristics. (For details, see ch. 7.)

The Statistical Methods Division (SMD) provided the sample weighting procedures that were applied by computer to each area in three successive stages, within which ratio estimation occurred. For persons, the first stage had 17 household-type groups; the second stage, just 2 groups—householders and nonhouseholders; and the third stage, as many as 160 age sex-race-Spanish origin groups. Ratio estimation for housing units followed a similar pattern, except that it was done in two stages for occupied units and in one for vacant units: The first stage for occupied units employed 16 household-type categories and the second, as many as 190 tenure-race-Spanish origin-value/rent groups. For vacant units, there were just three groups.

In addition to the basic long-form sampling scheme, i.e., 1-in-2 for incorporated places with fewer than 2,500 persons (based on precensus estimates) and 1-in-6 elsewhere, there were two subsamples—one in which every other sample questionnaire did not undergo coding for place of work and migration and the other, one of enumeration districts to provide the Early National Sample. Changes were made to these procedures following test runs of data from Delaware, Montana, and Rhode Island.

Analyzers

In addition to the ED-level data checks through diary review, followed by the sample weighting, the data for such areas as States, SMSA's, counties, minor civil divisions, and places were checked through elaborate computer programs called "analyzers." Complete-count data were checked with an analyzer generated from STF 1; the analyzer for sample data came from STF 3. Their purpose was to assure that the statistics for the larger areas conformed to expected levels based on the previous census or intercensal surveys. The analyzers also were used to check population and housing-unit count totals and weights used in sample products.

Confidentiality and Suppression

Census data were refined through many processes, but no figure was released without a final check. All data products—special and general—passed through a variety of analyses to

ensure that the statistics had been tabulated properly and that no confidential information would be revealed. To maintain the confidentiality promised to respondents and required by law, the Census Bureau withheld, or "suppressed," tabulations of characteristics of very small groups of people or housing units. In printed and microfiche reports, each suppressed data item was replaced by ellipses (...); on summary tape files, special "flags" denoted suppressed data.

Certain basic counts were not suppressed, even if an area had a count of only one. All other data might be suppressed under certain conditions (see below), primarily where the size of the population being characterized was less than a specified threshold. The suppression criteria differed for population data and housing data, and the thresholds were higher for sample estimation than for complete counts. The application of these thresholds resulted in what is known as "primary suppression." In addition, "complementary suppression" was applied to avoid the possibility of identifying an individual person, household, or housing unit by subtraction. Originally, complementary suppression was to extend to detailed race groups, among Spanishorigin types, between Spanish origin and not Spanish origin, or among the race-by-Spanish-origin/not-Spanish-origin groups. In December 1981, it was decided that in these cases, the suppression would be unnecessary and be a disservice to users. A set of revised rules, which are summarized below, was adopted, and applied to summary tape file (STF) 2 and its products, as well as to all subsequent STF's, tabulations, and products that used 1980 census detail files. However, data extractions or summaries of STF 1 that required suppression continued to use the original rules and/or indicators present in those files. Complementary suppression also was used if the number of units of one type of tenure (owner, renter) failed to meet a required minimum. In such cases, the data for housing characteristics were suppressed for both owner- and renteroccupied units.

The suppression rules were as follows:

- The following population counts will never be suppressed—total, race counts (detailed groups and totals for American Indian, Eskimo, and Aleut; for Asian and Pacific Islander; and for any other groupings), Spanish origin counts (total and four types and any group of types), Spanish origin by race (or race group) counts, and not of Spanish origin by race (or race group) counts.
- The following housing counts will never be suppressed—total housing units, seasonal/migratory housing units, total year-round housing units, vacant year-round housing units, and occupied housing units by race and Spanish origin, as described in (1) above.
- 3. 100-percent population characteristics based on persons for the groups specified in (1) above will be suppressed if there are 1 to 14 persons in the group; sample characteristics will be suppressed if the weighted count is 1 to 29.

- 4. Population characteristics based on families or households will be suppressed if there are 1 to 4 (sample, 1 to 9) occupied housing units with a householder of the specified race/Spanish origin group.
- Housing characteristics will be suppressed if there are 1 to 4 (sample, 1 to 9) housing units in the critical universe (e.g., occupied housing units, renter-occupied housing units).
- There will be complementary suppression between owners and renters and between summary or "major" race groups (i.e., White; Black; American Indian, Eskimo, and Aleut; Asian and Pacific Islander; and Other).
- 7. There will not be complementary suppression between detailed race groups, Spanish origin types, Spanish/not Spanish total, or race by Spanish/not Spanish groups; nor across geographic areas within a file or between files.

Data Review

To ensure that the edit specification programs (see ch. 12) produced reasonable results, the Bureau's professional staff reviewed the data summaries prepared from the computer tapes. This might require as many as 27 subject-matter specialists working from 2 hours to 3 weeks on a given summary, depending on the number of problems encountered, the time allowed, and the geographic level of detail at which the data were tabulated.

The Population (POP) and Housing (HOUS) Divisions were the principal Bureau units involved in this process; as many as 30 people in their various branches reviewed files and tables. A key staff member in each of the two divisions was responsible for coordinating the materials, answering the reviewers' questions, and deciding on the review time to be allowed. Usually, the reviewers immediately notified the computer programmers about the problems they found and thus tried to determine why the problems occurred. Later, the Decennial Census Division (DCD) received a summarization of all the known and potential problems/errors.

The subject-matter divisions (POP and HOUS) gave the decennial-area programming staff very detailed edit specifications covering how the various census questions should be handled during the editing process. These specifications were supposed to cover all possible combinations of responses. To determine if the edit programs were reasonable, logical, and written correctly, specialists reviewed the detail-tabulation files (generated on the basis of the specifications) at the State level, and the tabulation and report files at every level of geography (such as the county, census tract, and block). These materials were prepared on a flow basis by State. The detail files, unlike the other aggregate data records, contained only the unit records for each individual, and hence a detail file was a unit record instead of a data summary.

The reviewers had to pay special attention to allocation, i.e., supplying information where none was reported. Population edit specifications allowed the computer to go anywhere in the detail file for a given area to find a reported person with similar characteristics to impute in place of missing answers. Housing imputations were based on geographic proximity; where data were incomplete, the computer went back to the previously processed housing unit with reported characteristics, as experience had indicated that adjacent housing units usually were similar. In those rare instances where the first housing-unit record in a file had missing data, an anticipated average was imputed.

The edit-specification programs were based on specialists' judgment of what the data would look like (i.e., how often a particular question would meet with nonresponse and what should be done if it did). The review determined if these judgments were correct or if the edit specifications created a bias.

Specialists checked the data for three test States (Rhode Island, Vermont, and Montana) in greater detail than for the other States, under the assumption that problems associated with programming errors could be identified and corrected for the test States.

The headquarters staff completed its editing and reviews of the individual record files for the 100-percent data in January 1981 and the sample data in June 1982, respectively. It was only after the computer edits, sample weighting, and review were completed that edited information about persons and housing units, together with their associated geography could be stored on the basic record tapes (BRT's) from which all 100-percent and sample tabulations were produced.

When errors were found after the review, errata sheets were issued to accompany the published reports rather than alter the detail files. No attempt was made to generate products from edit tallies or analyzer programs.

TABULATION

Once the above processes were completed, edited data about individuals and housing units, together with associated geographic information, were stored on basic record tapes (BRT's). All 100-percent and sample tabulations were made from these tapes. Although the BRT's did not contain names and addresses, they did have detailed geographic codes and household data that could result in the disclosure of data for individuals; therefore, these tapes were confidential and could be used only by Bureau employees preparing statistical products.

BRT's were developed separately for 100-percent and sample data, and were processed both by type and by State. For review and tabulation purposes, there was a 100-percent edited detail file (EDF) and a sample edited detail file (SEDF) for each State. An EDF (either 100-percent or sample) contained a (tabulation) geographic header, followed by all corresponding

units and their housing characteristics. Each unit had associated with it the resident persons and their characteristics (person packets). An SEDF also contained weights for each sample unit and person. The EDF and SEDF BRT's were reviewed and cleared prior to tabulation.

Data summaries were prepared on computer tapes from the BRT's for all areas, including census blocks, census tracts, places, and counties. The resulting internal summary tapes

containing these tabulations were the source for tabulations that appeared in the printed reports, microfiche, and summary tapes made available for public use. The EDF BRT's also were used to prepare public-use microdata samples (PUMS), and both BRT's and summary tapes were used to produce special tabulations at user request and expense. These products and how they were produced are discussed in chapter 8.